



Bachelor's Thesis

Relationship between inbreeding coefficients, body weight, and litter size for 4 Danish dog breeds

By: Klara Hanefeld Egeriis, qgd922



Figure 1: Berner Sennen Klubben. (n.d.).

Academic Advisor: Bernt Guldbrandtsen

In Collaboration with: Dansk Kennel Klub

Date: June 12th, 15 ECTS

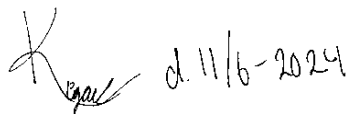
Acknowledgement

This academic research is prepared by Klara Hanefeld Egeriis as the 15 ECTS Bachelor written assignment in block four, year three of Animal Science at The Faculty of Science, University of Copenhagen. The objective of the research is to investigate the research question in a matter to enlighten other students as well as the members of Dansk Kennel Klub and others to which this work seems relevant.

The project is developed in collaboration with Bernt Guldbrandtsen and Helle Friis Proschowsky. A sincere thanks to Bernt Guldbrandtsen, my academic advisor, for his invaluable assistance with statistical analyses and R programming. A heartfelt thank you to Helle Friis Proschowsky for providing the pedigrees of the four dog breeds and for her ongoing communication with the Dansk Kennel Klub. Her role in enabling this project and sharing sensitive information has been invaluable.

Furthermore, a thanks to Laila Lehmann for her patience to locate the necessary data for this research which was led by diverse challenges.

Submission date: June 12, 2024



Signature: _____

Abstract

The purpose of the thesis was to investigate the relationship between inbreeding, body weight and litter size of four different dog breeds. This is done by analyzing existing studies, results based on own data and a discussion of this based on the theoretical framework. This quantitative study investigates relationships between explanatory variables and the independent variable in a statistical manner using linear models. The empirical data is based on pedigree data provided by the Dansk Kennel Klub. The theoretical framework is based on the book *Kompendium I Almen Husdyravl* by Bernt Guldbrandtsen, which has provided insight into aspects of inbreeding, genetics and populations. Otherwise, the statistical method is described based on *Statistical Data Analysis for the Life Sciences* used in the Statistical data course of this education. The results provided insight that the Rottweiler was the breed with the highest inbreeding coefficient ($F_{it} = 1.3$), but that inbreeding rates were generally low $< 1\%$. Kleiner Münsterländer had the largest average litter size, while Berner Sennen delivered the lowest. The concluding remarks imply that litter size for all four breeds is significantly negatively affected by maternal age, but the medium breeds are also affected by parity. The Rottweiler was the only breed significantly affected by maternal inbreeding. This project further aimed to fill the gap of lacking research on the specific research question by providing a focused analysis on the selected four breeds.

Resumé

Denne afhandlings formål var at undersøge sammenhængen mellem indavl, kropsvægt og kuldstørrelse hos fire forskellige hunderacer. Dette ved en analyse af eksisterende studier på området, resultater baseret på egne data og en diskussion af dette på baggrund af den teoretiske ramme.

Dette kvantitative studie undersøger forholdet mellem forklarende variable og den uafhængige variabel statistisk ved hjælp af lineære modeller. De empiriske data er baseret på stamtavledata leveret af Dansk Kennel Klub. Den teoretiske ramme baserer sig primært på bogen *Kompendium i Almen Husdyravl* af Bernt Guldbrandtsen, som har givet indsigt i aspekter af indavl, genetik og populationer. Rammen indeholder også en beskrivelse af den statistiske metode med udgangspunkt i *Statistical Data Analysis for the Life Sciences*, bogen der anvendes i kurset Statistisk data på denne uddannelse. Resultaterne gav indsigt i at Rottweileren var racen med den højeste indavlscoefficient ($F_{it} = 1.3$), men at indavlsgraden generelt var lav $< 1\%$. Kleiner Münsterländer havde den største gennemsnitlige kuldstørrelse, mens Berner Sennen leverede den laveste. De konkluderende bemærkninger indebærer, at kuldstørrelsen for alle fire racer er negativt signifikant påvirket af moderens alder, mens de mellemstore racer også er påvirket af paritetsnummer. Rottweileren var ydermere signifikant påvirket af moderens indavlsgrad. Dette projekt sigtede også mod at udfylde et hul i den eksisterende forskning ved at levere en fokuseret analyse af de udvalgte fire racer og deres specifikke forhold mellem indavl, kropsvægt og kuldstørrelse.

Content

1. INTRODUCTION	7
2. THEORETICAL FRAME	8
2.1 Breeding.....	8
2.1.1 Breeding Practices	8
2.2 Genome	9
2.2.1 Genetic Variation	9
2.2.1.1 Loss of Genetic Variation	10
2.2.2 Genetic Drift	10
2.2.3 Complex Traits.....	10
2.3 Populations	11
2.3.1 Effective Population Size.....	11
2.4 Inbreeding	11
2.4.1 Inbreeding Coefficient	11
2.4.2 Inbreeding Depression	12
2.4.3 Wright's F-statistics	13
2.5 Statistical Methods.....	13
2.5.1 Linear Models	14
2.5.1.1 Fitting the Regression Line	14
2.5.1.2 Estimation of Least Squares.....	14
2.5.2 Multiple Linear Models	15
2.5.3 Normal Distribution	15
2.6 R	16
3. DESIGN AND METHODS	17
3.1 Search Strategy	17
3.1.1 Limitations	17
3.2 Statistical Analysis	18
3.2.1 Data Analysis	20
3.2.1.1 Statistical Analysis.....	20
3.2.1.2 Data Definitions	21

3.2.2 Limitations	22
4. ANALYSIS	23
4.1 Effects on Litter Size	23
4.1.1 Inbreeding and Genetic Diversity	23
4.1.2 Body Size and Age of Parents.....	25
4.2 Chapter Conclusion	26
5. RESULTS.....	27
5.1 Distribution of Litter Size	27
5.2 Pedigree Completeness	28
5.3 Numbers in Pedigree	29
5.4 Inbreeding Trends	30
5.4.1 Wrights Coefficients	31
5.4.1.1 Bernese Mountain Dog	32
5.4.1.2 Border Collie.....	32
5.4.1.3 Kleiner Münsterländer	33
5.4.1.4 Rottweiler.....	34
5.5 Statistical Analysis	35
6. DISCUSSION	39
6.1 Effects on Litter Size	39
6.2 Inbreeding Trends	42
6.3 Distribution of Litter Size	44
6.4 Methods	45
6.5 Challenges in This Study	46
6.5.1 Scientific Uncertainty	46
6.5.2 Confounding and Causality.....	46
7. CONCLUSION.....	48

LITERATURE	49
APPENDIX	52
Appendix 1 – Breed information from DKK.....	52
Appendix 2 – R-coding, linear regression.....	53
Appendix 3 – Distribution of litter size for Bernese Mountain Dog	54
Appendix 4 - Distribution of litter size for Border Collie.....	55
Appendix 5 - Distribution of litter size for Kleiner Münsterländer	56
Appendix 6 - Distribution of litter size for Rottweiler	57

1. Introduction

The dog, *Canis familiaris* is the earliest domesticated animal (~35,000 BP) and descends from the wolf, *Canis lupus* (Galibert et al., 2011). The tameness begins according to early archaeological investigations in the Northern part of China about 300,000 years BP. Today, the population consists of several hundred breeds, each developed for specific purposes such as herding, guarding, hunting, or serving as companion animals, due to selective pressure and intensive breeding practices (Galibert et al., 2011). Today in Denmark Dansk Kennel Klub (DKK) is one of the oldest associations, and functions as an organization which purpose is to foster interest in healthy purebred dogs and hereby maintain Danish Dog Pedigrees, advice in breeding with mentally and physically healthy dogs in collaboration with experts, and much more. So far DKK is the only Danish organization recognized by the international federation of national kennel clubs, Fédération Cynologique Internationale (FCI), an association of national kennel clubs (Dansk Kennel Klub, n.d.-c). The club has established ethical recommendations for the breeding: *“including age of the female dog at the time of mating to responsibility for hereditary disorders, as well as the socialization and imprinting of the puppies”* (Dansk Kennel Klub, n.d.-a.). It is in the breeder's interest to breed large litter sizes which leads to the purpose of this assignment which aims to analyze the potential associations between inbreeding coefficients, body size, and litter size of four dog breeds with two different body sizes, based on pedigrees from DKK. Understanding how inbreeding affects litter size can help improving genetic health and longevity due to the negative effect of inbreeding on genetic diversity, resulting in higher incidences of hereditary diseases and reduced overall fitness (Calboli et al., 2008). With further knowledge due to the aim of this research can help provide breeders develop strategies to mitigate negative effects due to inbreeding. The following questions are to be answered:

How does inbreeding of the Bernese Mountain dog, Border Collie, Kleiner Münsterländer and the Rottweiler with body sizes medium and large affect litter size?

- Which inbred trends are seen for the four breeds?
- What factors affect litter size?
- Is there any evidence that body size or breed affect litter size?

Finally, the results will be published in a popular format in DKK's magazine to inform and educate dog breeders and owners about the potential consequences of these findings and to share the research more broadly.

2. Theoretical Frame

This section aims to clarify the theoretical framework underlying the scope of the project. Firstly, a general presentation of breeding and practices herewith. Next a presentation of the background of the genome, leading to genetic variation, genetic drift and complex traits. A presentation of a population and the effective population is presented and finally the quantitative statistical background of the analysis is mentioned.

2.1 Breeding

Breeding is a practice where human beings have taken control of which individuals should be mated by a process for selection of individuals due to different traits. The aim differs from the breeds, yet the objective is to produce offspring with the desired characteristics (Guldbrandtsen, 2023, p. 1). Random mating is on the other hand, mating would happen between two random individuals regardless of their genotypes (Guldbrandtsen, 2023, p. 48).

2.1.1 Breeding Practices

The commonly used breeding practice is selective breeding, where the methods rely on quantitative genetic principles, and sometimes by including molecular genetic information (Hill, 2017). To ensure effective results during the selective breeding, the population must include genetic variation to avoid potential inbreeding depression. A breeding program consists of different elements visualized in figure 2:

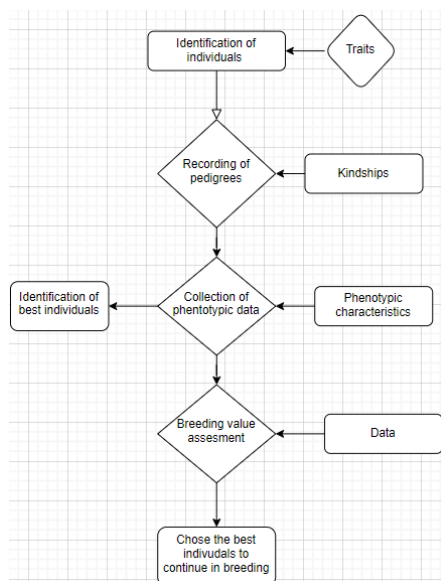


Figure 2: Elements of a breeding program to identify individuals to continue in breeding.

The selection of individuals for breeding can be more subjective in the case of dogs. This narrow on specific characteristics can result in lower fitness and welfare for some dog breeds due to the intensive breeding practices (Hill, 2017).

2.2 Genome

Gene is the physical unit transmitted from parent to offspring and influences hereditary traits. An individual genotype consists of a set of genes, whereas the physical response to a genotype is the phenotype. Most traits are affected by more genes, especially complex traits such as haircolor, weight and reproductive fitness, and some is affected by the environment, why one genotype can result in more phenotypes. Those different traits are affected by alleles, which defines the trait from that gene. One gene consists of two alleles, and one gene is the result of heritability from both parents, whereas one allele comes from the mother and the other from the father. Those alleles can be homozygote meaning identical, or heterozygote meaning different (Hartl, 2020a). In the ultimately population, all copies of alleles are from earlier ancestors. In order to determine the relationship between alleles in an individual and those in their ancestors or previous generations, it is necessary to identify the parents of the individual in question (Guldbrandtsen, 2023, p. 85). This is typically done by pedigrees, to understand the genetic background of the individual. Alleles can be identical by descent (IBD), meaning either direct copies from parents to offspring or if an allele can be tracked to one of the parents alleles in earlier generations. Alleles that are identical by descent are also always identical by state (IBS). However, this is not always the case in reverse, this results in alleles with same shapes or sequence visualized in phenotypic lookalikes (Guldbrandtsen, 2023, p. 86).

2.2.1 Genetic Variation

Genetic variation refers to the variation in the sequences of the organisms DNA that contribute to the phenotypic traits (Barton & Keightley, 2002). It is interesting because it is prerequisite for breeding progress. As relatedness increases the genetic variation must decrease and that will affect the heritability, as seen in the formula for breeding progress:

$$\Delta G = \frac{ir_{\hat{A}A}\sqrt{h^2}\sigma_P}{TG} = \frac{ir_{\hat{A}A}\sigma_A}{TG}$$

Whereas ΔG is the expected breeding progress per unit of time, i is the intensity of selection, h^2 is the heritability in narrow sense, σ_P is the phenotypic spread and TG is the generation interval (Guldbrandtsen, 2023, p. 160). The heritability plays a role in the formula for ΔG , so when the genetic variation decreases, the heritability decreases as well which result in lower breeding progress

(Guldbrandtsen, 2023, p. 126). The cause of genetic variation can be due to increased relatedness (genetic drift) or selection, which is often seen in the breeding of dogs. Selection transforms genetic variation within a population into desired traits leading to changes in the populations mean value affecting breeding progress.

2.2.1.1 Loss of Genetic Variation

Genetic drift, described in the next section, leads to loss of genetic variation as fewer and more related individuals would be a part of the breeding in the populations. It affects the variation and all possible traits no matter selective or not. Over time, this leads to a decrease in the variation of all traits (Guldbrandtsen, 2023, p. 165). Selection affects the variation as it aims to increase the frequencies of alleles with positive effects on traits wished for in the breeding goal (Guldbrandtsen, 2023, p. 166).

2.2.2 Genetic Drift

Genetic drift is random changes in gene frequencies in a population over generations (Guldbrandtsen, 2023, p. 103). It affects the alleles in the population independently on their influence on reproduction and survival and over time the genetic composition becomes increasingly distinct from their starting point of ancestors. Genetic drift can lead to a reduction in genetic variation, as all populations are final. (Guldbrandtsen, 2023, p. 104). Furthermore, it restricts the possibility for future selection including for traits that currently have no value as the additive genetic variance is lost (the variance of breeding values in the population) (Guldbrandtsen, 2023, p. 126, 73).

2.2.3 Complex Traits

Complex traits are traits affected by effects of multiple genes (polygenes) and environmental factors. Examples of so includes weight, body composition, hair – and eye colour. Classifying; *Metric traits* as measured on scale (e.g., height and weight) and *Categorical* as counted such as litter size (Hartl, 2020c). Heritability is the proportion of phenotypic variance that can be attributed to genetic variance in broad sense (H^2) and narrow sense (h^2). Broad sense measure to total variance contribution to the variance in the trait including all genetic variance components whereas narrow sense is focusing on additive genetic variance (Hartl, 2020c). Interactions between genotype and environment occur when the effect of a genotype varies across different environments, which complicates the prediction of phenotypic outcomes (Hartl, 2020c).

2.3 Populations

Here a population is a group of individuals of the same species sharing the same pool of genes. The effective population size is a measure of how quickly a population is losing the genetic variation and depends on the number of individuals available in the breeding. Different factors affect the effective population, such as the number of animals included in the breeding population. The ratio of sex affects the effective population since the highest possible population of so is with an even sex ratio. The number of offspring given as variance since a decrease in offspring leads to lower genetic variation and possible increase the inbreeding of the populations. The effective population size can change and is sometimes affected by a bottleneck due to environment changes (Guldbrandtsen, 2023, p. 108-113).

2.3.1 Effective Population Size

The rate of genetic drift can be measured by the effective population size N_e , based on molecular data or pedigrees. The effective population size is a measure of the rate a population is currently losing genetic variation. The effective population size depends solely on current breeding practices and the number of available breeding candidates. The strength of the genetic drift is affected by different factors such as number of animals, sex ratio, variation in number of offspring and selection. Moreover, fluctuations over time is affective, and a period of bottleneck can affect the effective population size. Breeding dogs, is often based on selection with clear selection criteria, which typically results in breeding with fewer males and more closely related individuals, hence lower effective population sizes (Guldbrandtsen, 2023, p. 108-113).

2.4 Inbreeding

Inbreeding is a result of two identical alleles in a gene inherited from ancestors in the generations (Guldbrandtsen, 2023, p. 85). The reasons for inbreeding are mating between related individuals or a result of genetic drift due to the ultimately population size.

2.4.1 Inbreeding Coefficient

The only way to conclude whether a genotype is ad result if inbreeding, is by calculating the probability of so based on pedigrees. The probability of inbreeding is called *the inbreeding coefficient*, F (Guldbrandtsen, 2023, p. 87). The principle for calculating inbreeding is the loop-method (Guldbrandtsen, 2023, p. 93). The pedigree consists of lines representing gametes contributed from both parents, and the figures identify individuals in the population according to sex (figure 2). Firstly,

to calculate F_I ancestors must be identified, here, A, and the lines must be followed from I to A to find all possible loops, here one. From that the contribution of loops is calculated by the number of meiosis (lines) in every loop from parent to ancestor as:

$$\text{Contribution}_A = \left(\frac{1}{2}\right)^{n+n'+1}$$

Whereas n is the loop from mother and n' from father. The possible contributions are to be added together to find F_I (Guldbrandtsen, 2023; Hartl, 2020b). In general:

$$F = \sum_i \left(\frac{1}{2}\right)^{n_i+n'_i+1} (1 + F_{A,i})$$

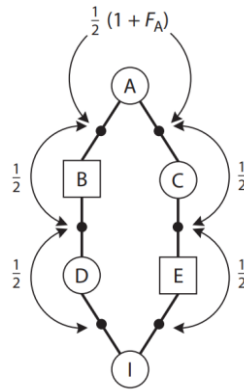


Figure 3: Principle for identifying loops. Numbers representing the probability of IBD (Hartl, 2020b).

2.4.2 Inbreeding Depression

Inbreeding depression is the cause of inbreeding where the individual's traits deteriorate, and affects all traits, but somehow traits with low heritability are more affected and those with high heritability (Guldbrandtsen, 2023, p. 98). One hypothesis on occurrence of inbreeding depression is that inbreeding lower the frequency of heterozygosity. Another hypothesis on the occurrence of inbreeding depression is the result of higher frequency of recessive deleterious alleles, unfavourable for the individual (Guldbrandtsen, 2023, p. 98).

The history of taming, domestication to selective breeding for the humans to take control over different traits in animals offspring, often leads to inbreeding depression due to the ultimately population size. Inbreed individuals in the population often contributes to unfavourable traits. Even so, the inbreeding can cause effects on the offspring, inbreeding is not heritable, and will stop by mating an inbreeding individual to a non-inbreed animal. In dogs the selective breeding often leads to challenges for the individual, why pedigrees and inbreeding coefficients must be considered (Guldbrandtsen, 2023, p. 100).

2.4.3 Wright's F-statistics

Wright's F-statistics are used in population – and evolution genetics to understand genetic variation between and among populations to analyse selection (Holsinger & Weir, 2009). The coefficients include:

The total inbreeding coefficient F_{it} measures the inbreeding of the whole population by taken intra – and inter subpopulations genetic variation into account (genotypes). A high value contributes to highly inbreeding in the population as the consequence of deficit of heterozygotes (Holsinger & Weir, 2009). A negative value or equal zero indicated genetic health in the population with random mating and low inbreeding. In the long run, F_{it} gives an overall view of the genetic health of the animals and can contribute to long run breeding strategies to secure genetic diversity (Holsinger & Weir, 2009).

The fixation index F_{st} due to genetic drift calculates the genetic variance among subpopulations, a measure of how much of the total genetic variation that is due to difference in allele frequencies. A high value indicates remarkable genetic variance among subpopulations, by to the variation of allele frequencies. If the value is equal to zero, no genetic variance occurs. A negative value is normally due to error in data or analysis since it measures variance, which per definition cannot be negative (Holsinger & Weir, 2009).

The inbreeding coefficient F_{is} due to non-random mating calculate the inbreeding among a subpopulation by comparing the frequencies to Hardy-Weinberg to indicate a positive or negative number of heterozygotes. A high value indicates a deficient of heterozygotes which equals high inbreeding. If it is equal to zero, the Hardy-Weinberg is met and leads to random mating without inbreeding. A negative value shows a surplus of heterozygotes (Holsinger & Weir, 2009).

2.5 Statistical Methods

Statistics in matter of animal genetics is about *“making statements about a population from data observed from a representative sample of the population.”* (Ekstrøm et al., 2015 p. 1). In this assignment the population includes data of litters for the chosen breeds, which is a sample of the actual data, since it is a part of the population. Statistics can be used to describe parameters of the sample size of the population, here, inbreeding, litter size and body size of the breed. The mean value is often used to describe a trend to the sample size, which somehow can be a description of the population (Ekstrøm et al., 2015 p. 1).

2.5.1 Linear Models

It is often relevant to describe a variable as a function of another, and can sometimes rely on well-known hypothesis, whereas other times no prior knowledge is available, but data can identify relationships empirically. The linear equation can be written:

$$y = \alpha + \beta * x$$

Alfa function as the intercept, and the value of y whenever x equal zero. Beta refers to the change in y for each unit change in x and is called the slope. When the model is visualized y refers to the dependent viable while x is the explanatory (Ekstrøm et al., 2015 p. 27).

Even though most real-life data rarely exhibit a clear linear relationship between variables, this model may still provide the best fit for the data.

2.5.1.1 Fitting the Regression Line

The purpose of fitting the regression line is to identify the “best possible line” to describe the data. Estimates of the values alfa and beta can predict for any x the y-values for the model:

$$\hat{y}_i = \hat{\alpha} + \hat{\beta} * x_i$$

Where the hat above alfa and beta indicates unknown values of so, but functions as estimates. By the predicted values, it is possible to evaluate how well the model fits to the true observed data by looking at the residuals defines as:

$$r_i = y_i - \hat{y}_i$$

Residuals functions to measure how far away from the expected values the true observed are to the regression model, here as a straight linear model. Small values indicate a close relationship between expected and observed values with respect to the chosen model (Ekstrøm et al., 2015).

2.5.1.2 Estimation of Least Squares

The sum of residuals is not giving an adequate evaluation of the fitting model, since a large positive residual can be canceled by a corresponding negative one. The method of least squares includes squared residuals before adding those to the model to prevent the outliers to affect each other and by that optimizing the residuals by minimize the sum of the squared residuals. By that, the total deviation from the regression line is as small as possible:

$$\sum_{i=1}^n (y_i - \alpha - \beta * x_i)^2$$

2.5.2 Multiple Linear Models

In multiple linear regression it is assumed that more than one explanatory variable x is observational for each unit. The model can be defined by:

$$y_i = \alpha + \beta_1 x_{i1} + \dots + \beta_d x_{id} + e_i, \quad i = 1, \dots, n,$$

The residuals are assumed to be independent and normally distributed. Here, $\beta_1, \dots, \beta_d$ are partial slopes and α is the intercept as the expected value whereas $x_1 = \dots = x_d = 0$. An unit change in x_k corresponds to an expected change in y of β_k , if all other variables remain unchanged. The estimated is to be found likewise for the linear regression. When investigating a model like the multiple linear regression, a hypothesis is normally made whereas one explanatory variable equals zero (e.g., $H_0: \beta_d = 0$), which corresponds to no effect or influence on x_d on y given the other explanatory variable remain in the model (Ekstrøm et al., 2015, chapter 8).

2.5.3 Normal Distribution

The normal distribution functions to describe random variation in the dataset investigated. It visualizes the variation by the mean value of the wished investigated parameter, for example litter size in the sample size of the population of the Bernese Mountain dog. A histogram is normally used in such cases, where the mean values are put together in the middle and datapoints outliers to the sides by seeing the frequency of those. The histogram is normal distributed as soon as the sum of the areas of rectangles equals one. The area under the curve from a to b refers to the probability that a random observation falls within this exact interval (Ekstrøm et al., 2015, chapter 4).

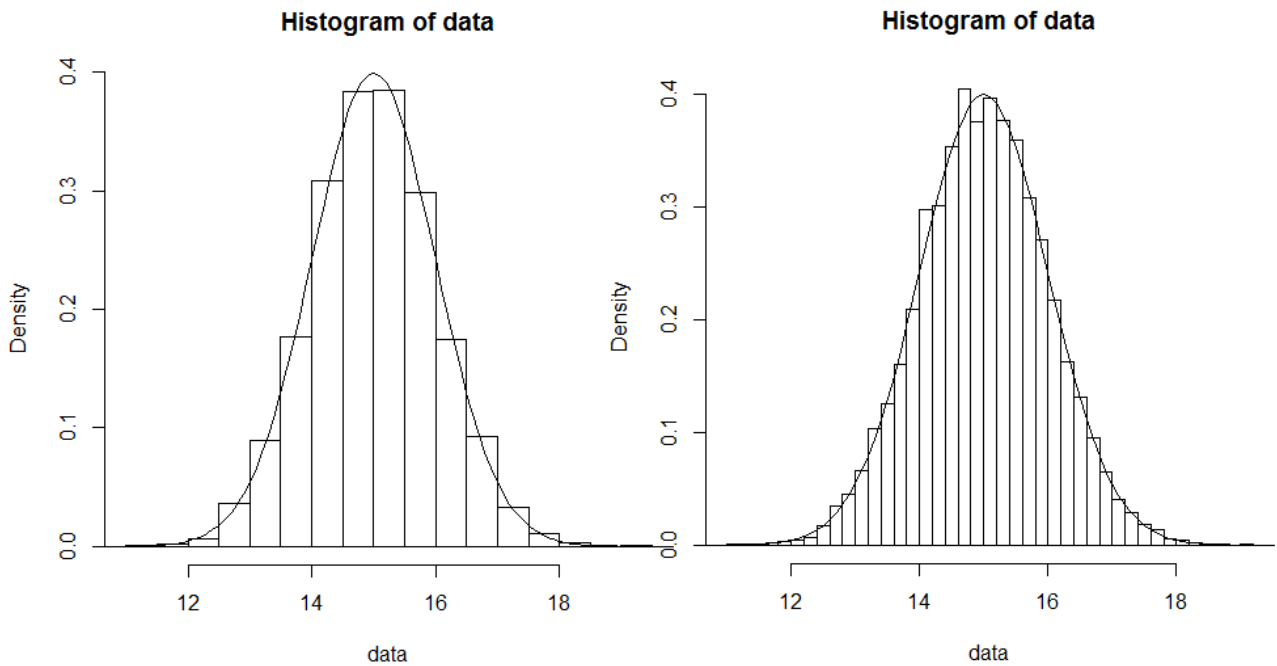


Figure 1: Visualize a histogram of the normal distribution by rectangles below the curve.

2.6 R

R is a versatile programming language and environment optimized for statistical analysis and graphic creation. It supports a wide array of statistical methods, including linear and nonlinear modeling, classical statistical tests, time-series analysis, classification, and clustering. Users can customize and expand its functionality by creating new functions, making R a favored platform for statistical research with its open-source accessibility. It features advanced graphical capabilities, allowing for the creation of publication-quality plots that include mathematical symbols and formulae. R ensures robust data manipulation, storage, and analytical tools, along with comprehensive facilities for handling arrays and matrices. Additionally, its extendability through packages allows for broad application in modern statistics. R operates under the GNU General Public License and is compatible with multiple operating systems including UNIX, Windows, and MacOS, showcasing its broad adaptability in the statistical computing field (R Core Team., 2024).

3. Design and Methods

In this section the method of the search strategy of literature for the analysis is explained and clarified as well as the data processing of the pedigrees delivered from DKK and the choices made to be used in the section of results. Moreover the data is described more detailed to clarify what is the basis of the analysis and to clearly document the conclusions made of so to reduce ambiguities.

3.1 Search Strategy

The search for information and relevant literature was done by a systematic search in the database “Web of Science”. The keywords and search were divided by three different searches, constructed as follows:

- Dog Breeds AND Inbreeding OR pedigree analysis AND body size OR litter size - producing 47 results
- Litter size AND dog breeds producing 253 - results

The search identified a total of 300 articles. A general overview was made from the titles of the articles which restricted the number to 31 to ensure the ones relevant to this topic. Afterwards the articles were sorted based on the abstract and conclusion according to specific inclusion and exclusion criteria (Table 1), which ended up with 10 relevant articles.

3.1.1 Limitations

The following inclusion and exclusion criteria for the papers were made to keep a focus on a few areas related to the research rather than a broader overview. The focus is on litter size for domesticated dog breeds. Even so other species could seem relevant according to heritabilities and by that establish correlations, the focus has here been on dog breeds to simplify the topic and research. Inbreeding is relevant to the research question, why that is indeed an important factor to incorporate with the aim to discuss similarities and differences between the analysis and the results in this rapport. Weight is a factor related to the body size related to the two different weight classes of dogs to be analyzed.

Table 1: Inclusion and exclusion criteria were used to screen for relevant literature.

<i>Inclusion criteria</i>	<i>Exclusion criteria</i>
The publication included dog breeds and no other species.	The publication specifically investigated cancer or other diseases.

Inbreeding is researched or considered in this publication.	The publication included nonrelevant breeding goals.
The publication included heritabilities related to weight or other relevant keywords.	Behavior traits were examined in the publication. Wild dogs were the starting point for the study.
Litter size is incorporated in some form in the article.	The publication included cloning or the use of frozen semen. Supplements were given throughout the study.

Furthermore, specific search was done as the rapport unfolded if needed.

3.2 Statistical Analysis

The research results are based on pedigrees provided by the DKK, containing data from 1989 onward. To select four dog breeds to move forward with, breed data from 258 dog breeds, including information such as litter size, were analyzed (see Appendix 1). From this data, I focused on the first 47 breeds due to the restricted length of this assignment. Using data from the DKK's website, the average weight of each breed was registered (Dansk Kennel Klub, n.d.-b). Subsequently, a log-log regression was performed on the average weight and litter size (see Appendix 2 for the R code).

General linear models were used to analyze relationships among body size and average litter size. The model used was:

$$L = a \times W^b,$$

where a and b were constants, L was litter size, and W was average body weight in the breed.

The average body weight was conducted from DKK's website "*Hundeleksikon*" where the search for each breed provided the average weight of the breed, leading to the body size of the breeds including medium and large body sizes (Dansk Kennel Klub, n.d.-b).

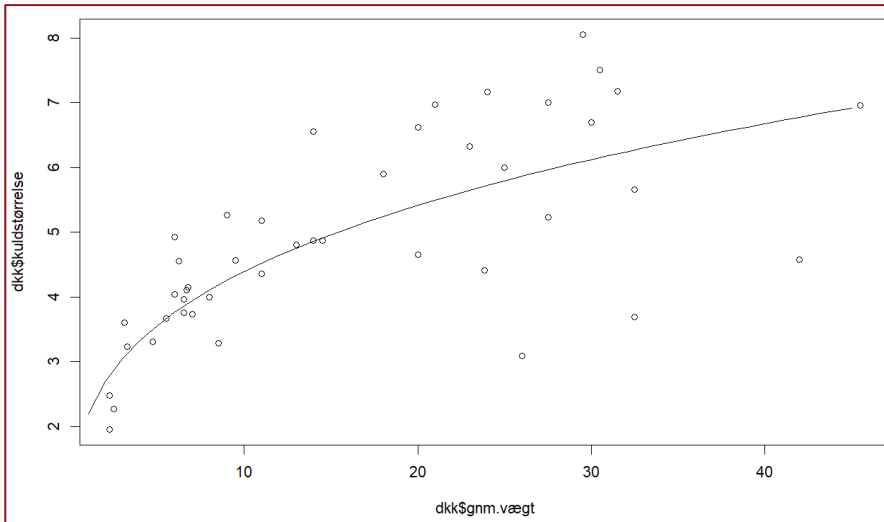


Figure 2: Linear regression over average weight and litter size for the 47 dog breeds delivered from DKK.

To visualize the linear correlation rather than the experimental function, log is taken to the factors:

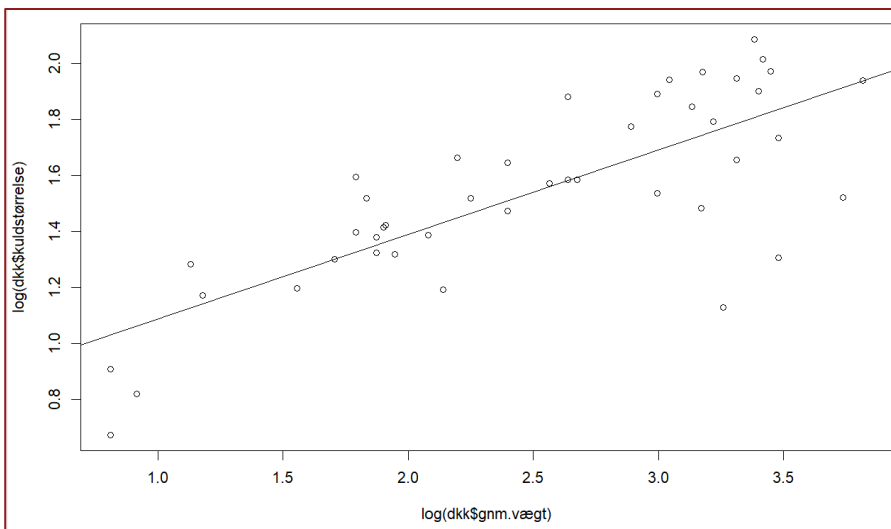


Figure 3: Log-linear regression over the average weight and litter size for the 47 dog breeds delivered from DKK.

From that, it was possible to see datapoint outliers and recognize those to work with further in combination with DKK's guidance to avoid dog breeds with genetic complications such as diseases that could affect the conclusion of the analysis. Breed selection criteria were:

1. Two breeds with medium body size, two breeds large body size.
2. Adequate information on litter size available from DKK.
3. Avoiding breeds with chondrodystrophy.

Four dog breeds were selected: Bernese Mountain dog, Border Collie, Kleiner Münsterländer and Rottweiler.

Ideally, we could have wished to make linear regression over all 258 dog breeds, instead of the first 47, and to work with breeds from three to four weight classes, instead of only medium and large dog breeds.

DKK provided pedigrees of the four dogs. The data were inspected. Several types of problems were resolved: Parents who were born after offspring, animals without their own entry in the pedigree, and animals appeared as both sire and dam. Cleaned pedigrees delivered by my supervisor. Subsequent analyzes were performed in R.

3.2.1 Data Analysis

Three different datasets were available: pedigrees over the four breeds, summary inbreeding statistics, and litter data.

Pedigree data included both Danish and foreign dogs.

Litter data consisted of litter number, sire, dam and their registration numbers, date of birth, mating (natural vs fresh semen vs frozen semen), birth type (natural vs caesarean section), number of male and female puppies born in the litter, number of male and female puppies that survived in the first three to five weeks till registration with DKK, inbreeding coefficient of dam, sire, and offspring, date of birth of the dam, and parity number. During analysis, the following data additional columns were created: year of birth, age of dam at birth of the litter, total numbers born (live plus stillborn), total number survivals (till age of three to five weeks), and fraction surviving (born alive). Data analyzed here was the number of puppies alive at the time of registration.

The dataset of **inbreeding coefficients** included year of birth, number born in the year, Wright's F statistics (Wright, 1950): F_{it} as the total inbreeding, F_{st} as inbreeding due to drift and F_{is} as inbreeding due to non-random mating, and average pedigree completeness. The years goes from 1980 to 2023. F_{it} was computed by custom scripts as the average realized pedigree inbreeding coefficient in each year-of-birth cohort. F_{st} was computed as the average pairwise coefficient of relationship among all dogs in each year-of-birth cohort excluding the relationship of a dog with itself. F_{is} was derived from F_{it} and F_{st} . Pedigree completeness was the average generations of pedigree-equivalents known counting the dog itself as 1.

3.2.1.1 Statistical Analysis

Linear regression was used to fit a linear model with the addition of effects putatively affecting the litter size. In this project, those effects were the effects of the inbreeding of sire, dam, and offspring, parity of the dam, year-of-birth of the litter, and age of the dam at birth of the litter. The purpose was

to investigate which of the aforementioned factors significantly affect litter size by progressively simplifying the model by removing factors without significant effects. For all breeds “year” was removed as first effect, next inbreeding of sire followed by the inbreeding of offspring. Finally, the inbreeding of the dam or parity was removed depending on the significant correlations. To simplify models, the explanatory power of two models was compared using analysis of variance (anova). Only if the more complex model had significantly greater explanatory power, was the more complex model preferred. The null-hypothesis (H0) is given:

H0: *Year has no effect on the litter size of the four dog breeds.*

H0: *Inbreeding of sire has no effect on the litter size of the four dog breeds.*

H0: *Inbreeding of offspring has no effect on the litter size of the four dog breeds.*

H0: *Inbreeding of dam has no effect on the litter size of the four dog breeds.*

H0: *Parity has no effect on the litter size of the four dog breeds.*

H0: *Age of dam has no effect on the litter size of the four dog breeds.*

The Wright’s F statistics was plotted as functions of year-of-birth. The distribution of litter size among the breeds was visualized by histograms and boxplots.

3.2.1.2 Data Definitions

In the following section, the meaning of the data columns will be reviewed.

Litter size

Litter size was defined as the number of puppies alive at the time the litter was registered with DKK by the breeder, typically around three to five weeks of age. Puppies, which died after registration were included in the litter size.

Age of the dam and parity

The age of the dam is defined as the age in days at time of birth of offspring of the litter in different parities. The parity is given from first to six. Parities of dams were computed as the number of each dam’s litter in the data in order of appearance. Some dams may have had litters recorded prior to the start of the data. Hence, the parities are not entirely correct for the first couple years. Data is only from the last ten years so the first parity for the dam in the first years, is not necessarily the first but could potentially be the third.

Inbreeding of dam, sire, and offspring

These are the inbreeding of mother to litter, the sire of the litter, and the inbreeding coefficient of the puppies in the litter.

Year

Year is given as the year the litter was born.

For all the statistical analysis, the program R is used described in the theoretical frame part 2.5.3.

3.2.2 Limitations

The research will not cover further breeds than the Bernese Mountain Dog, Border Collie, Kleiner Münsterlander or the Rottweiler. The effect of mating procedure or environmental effects is not investigated. The research is based on quantitative methods including the use of tools to calculate inbreeding coefficients based on common ancestors and employing linear regression, to investigate the correlation between explanatory factors such as inbreeding of dam and sire on litter size. Additionally it will examine the correlation between body size and litter size for the two body sizes studied (medium and large). The reason behind this limitation is to ensure a focused and feasible study within the timeframe and resources available. It is recognized that there are many factors that can influence litter size among different dog breeds, but this study will look exclusively at selected factors as variables. Previous studies on dog breeding have examined various factors affecting litter size, including genetic diversity, health conditions, and environmental influences with more. However, few studies have specifically analyzed the combined effect of inbreeding and body size on litter size within specific breeds. This research aims to fill this gap by providing a focused analysis on the selected breeds.

4. Analysis

The following section contains an analysis of the ten articles from the search strategy in order to give a reflection of what has been researched up to the present project for dog breeds in general. The section will be described based on themes selected and constructed to address the research question, which aims to investigate how inbreeding and body size affect the litter size of offspring for the four different breeds. The analysis is organized around the main theme of *effects on litter size*, with corresponding sub-themes described in the following section supported by studies.

Research including Leroy et al. (n.d.), McMillan et al. (2024) and Yordy et al. (2020) emphasizes the importance of understanding the influence of inbreeding and body size on litter sizes and the survival of the dogs. The following studies reveal intricate interplay between genes and environmental factors.

4.1 Effects on Litter Size

Litter size is investigated among studies to clarify or analyse what factors that could affect the size of offspring of dams, especially inbreeding (Borge et al., 2011; de Andrade et al., 2021; Kania-Gierdziewicz & Pałka, 2019; Leroy et al., n.d.).

4.1.1 Inbreeding and Genetic Diversity

Leroy et al., (n.d.) conducted a study on seven breeds, revealing a clear correlation between high levels of inbreeding and reduced litter size, as well as decreased survival rates of puppies up to two years of age ($p < 0.05$), both affected by inbreeding of the dam (estimation from -0.13 and -0.76) as well as the offspring (estimation from -0.27 to -0.65) which decreases the litter size. Supporting these findings, de Andrade et al. (2021) demonstrated that increased inbreeding depression led to reduced litter sizes in German Spitz dogs. The study indicated that inbreeding within the litter itself had a more substantial negative impact than the inbreeding of the dam with both to be significantly approved. Additionally, the study notes that a similar relationship has been observed in other breeds, including the Bernese Mountain dog, where higher inbreeding among male animals is correlated with litter size reduction. These results collectively emphasize that inbreeding negatively affects reproductive success by reducing viable gametes, as previously studied in felines. Overall, they conclude that there is a strong influence of puppies' inbreeding as well as an influence of dam breeding on litter size in the population investigated. de Andrade et al. (2021) reflects on the consequences for using linear regression to conduct a correlation, since it does not always occur linearly. Both studies based their investigation on data from dog kennels respectively the French Kennel club and Kennel Clube

do Rio Grande do Sul in Brazil. Their results indicate a future reduction in biological fitness due to inbreeding which affects the health and survival of the dogs (Leroy et al., n.d.; de Andrade et al., 2021).

Genetic diversity is recognized as a crucial factor for health and longevity, so research underscores the importance of maintaining genetic diversity to prevent inbreeding depression (Kraus et al., 2023; Yordy et al., 2020). Kraus et al. (2023) found that while the precise impact of genetic diversity on various outcomes is complex. Nonetheless, larger body sizes tend to be associated with shorter lifespans and increased age-related mortality. Conversely, these factors increase with higher genetic diversity.

Selective breeding practices, which can increase inbreeding, have notable effects on reproductive traits (Marelli et al., 2020). Marelli et al. (2020) examined the correlation between inbreeding and reproductive outcomes, finding that higher inbreeding coefficients are linked to increased occurrences of stillborn puppies and reduced litter sizes. Specifically, a 1% increase in the inbreeding coefficient of dams was correlated with a 0.06% increase in dead puppies per litter. A one percent increase in the inbreeding coefficient of the sires resulted in a 0.007% increase in the number of dead puppies. This illustrates how the dams' inbreeding level impacts the reproductive trait regarding litter size and litter composition (Marelli et al., 2020). The observed effect by both dams and sires, indicates that inbreeding at any level of the pedigree can negatively impact reproductive success.

Further, Kania-Gierdziewicz and Pałka (2019) investigated reproductive traits in five different dog breeds, focusing on litter size, sex ratio, and the number of male and female puppies. Their study found that inbreeding levels in parents could influence these traits, only with significant effects observed in the Tatra Shepherd Dog (TSD). Inbred dams were associated with increased litter sizes ($0.23, p < 0.05$), while inbred sires negatively impacted the sex ratio, resulting in more female puppies (0.20 estimation for females, -0.05 for males, $p < 0.01$). The correlation coefficients between litter inbreeding and litter size for the majority of the examined dog breeds were positive but not significant ($p > 0.05$) as well as the effect of the inbreeding of parents on fertility traits (litter size, number of male offspring in litter, number of female offspring in litter, sex ratio and sex difference respectively LS, NM, NF, SR and RN) was not significant ($p > 0.05$) (Kania-Gierdziewicz & Pałka, 2019). The method was based on data from breeders given to the Krakow department in Poland from 2000 to 2007 including data from 436 litters including 2560 puppies, with a sex ratio of 1303 males:1206 females and parents 163 sires: 228 dams.

4.1.2 Body Size and Age of Parents

The breed and size of a dog are factors that have been suggested to affect litter size (Borge et al., 2011). Yordy et al. (2020) found a negative correlation between body size and longevity, noting that larger dogs tend to have higher inbreeding coefficients. The study indicates that, while larger dogs generally produce larger litter sizes, they also tend to have shorter life expectancies. Kraus et al. (2023) supplement Yordy et al. (2020) by emphasizing that while larger dogs tend to have larger litters, their reproductive life expectancy is mainly influenced by body size rather than inbreeding depression. This suggest a complex interaction between genetics and physiological traits. However, genetic diversity was not significantly correlated with body size as observed in Yordy et al. (2020). Kraus et al. (2023) also revealed a significant negative correlation between body size and longevity, and a positive correlation between longevity and heterozygosity. Meaning, larger dogs in general tends shorter life expectancies with increasing age, but dogs with higher genetic diversity (heterozygote) tends higher life expectancy.

To investigate factors affecting litter size, Borge et al. (2011) conducted a retrospective study of 224 breeds using data from the Norwegian Kennel Club from January 2006 to December 2007. The mean litter size was 5.4 (0.025) puppies, with the largest litters seen in Rhodesian Ridgebacks (8.9 puppies, ± 0.6) and the smallest in Toy Poodles and Pomeranians (2.4 puppies, ± 0.1). Linearly analysis of 10,810 litters using linear regression showed a significant increase in litter size with breed size. the mean value of litter size increased with the size of breed, and the relation seemed significant by the model. Average litter size in the study of the different breed sizes is given in table 2.

Table 2: Showing the average litter size for five different dog size groups from Borge et al. (2011)

Miniature	Small	Medium	Large	Giant
3.5 (± 0.04)	4.2 (± 0.03)	5.7 (± 0.04)	6.9 (± 0.05)	7.1 (± 0.13)

The initial model, a simple linear regression, did not reveal clear trends on how age affects average litter size among the different breed sizes. However, the multivariable analysis did. It showed a curved relationship between age and litter size, indicating a significant interaction between age and breed size (Borge et al., 2011). For small breeds, the youngest and oldest dams had smaller litter sizes compared to those in the middle age group. For larger breeds, giving birth at a young age did not generally reduce litter size, but the number of puppies decreased as the dams aged. The largest average litter size was observed in dams having their first litter (5.7 ± 0.04), and this value decreased with increasing litter number (parity), reaching a minimum of $4.2 (\pm 0.24)$ by the sixth litter (Borge et al.,

2011). Leroy et al., (n.d.) showed that litter size increases until the second or third litter for several breeds, including Bernese Mountain Dog (BMD), German Shepherd Dog (GSD), and West Highland White Terrier (WHW), then litter size decreases in subsequent litters. This pattern emphasizes the importance of considering both maternal age and parity when assessing litter size (Borge et al.). The medium breeds exhibited a decline in litter size at an earlier age, with a notable decrease observed within the first five years. However, when adjusted for the dam's age in the multivariable analysis, the negative effect of the number of litters on average litter size was not significant (Borge et al., 2011). Overall, litter size was significantly influenced by breed size, mating method, and age of the dam. Parity of the dam did not affect litter size, and the effect of inbreeding was not investigated (Borge et al., 2011).

Strandberg et al. (2023) investigated genetic trends and variations in body weight across several dog breeds. Their study highlighted that both environmental and genetic factors affect a dog's body weight. Using linear regression, they found that breed, sex, year, and age significantly influenced body weight ($P < 0.001$), while the parity of the dam was also significant ($P = 0.03$), and litter size was nearly significant ($P = 0.06$).

Yordy et al. (2020) shows a strong negative correlation between body size and lifespan in dogs but note that it is unclear whether the higher rate of aging in large dogs is due to body size or other associated factors. Additionally, they found that inbreeding levels tend to increase with body size ($r = 0.18$, $p = 0.02$), and while large breeds tend to have larger litters and higher inbreeding levels than smaller breeds, the reduced life expectancy of large dogs is driven by body size rather than inbreeding depression.

4.2 Chapter Conclusion

The analysis of the mentioned studies shows an indication that both inbreeding, genetic diversity, parity, age and breed size affects litter size. Inbreeding and inbreeding depression in general leads to a decrease in litter size and lower survival rates whereas large dog breeds, despite bigger litters, have the tendency to shorter lifespans and affects reproductive traits. Somehow the age of the dam seems to affect litter size among few of the studies, but the breed size corresponds differently to that effects. Nevertheless, the number of puppies decreased as the dams aged.

5. Results

This section will study the research question in statistical manner, whereas linear regression is used as a model to describe the possible factors including the inbreeding of litter, dam, age of mother and finally parity number. The theoretical background of the models is described in the theoretical framework.

Sample size and inbreeding coefficients varies among the four chosen breeds (Table 3). Rottweiler with the highest inbreeding coefficient. Kleiner Münsterländer with the highest number of survivals in litters (mean = 7.2). The Bernese Mountain Dog with the lowest average litter size and lowest survival. However, the standard error for the Bernese Mountain Dog as well as Rottweiler regarding average litter size is the highest with 3.0. The Border Collie shows to be the least inbred breed of the chosen. Number of litters include relatively small sample sizes.

Table 3: An overview of the four breeds includes the number of animals analyzed, the inbreeding coefficient (Fit) for the year 2022 due to the higher number of individuals registered, the average inbreeding among litters (both given in percent), the average litter size with standard error, and finally, the average number of surviving puppies in litters with standard error.

Breed and size	N dogs	N litters	Average F_{it} year 2022 [%]	Average litter inbreeding [%]	Average litter size [std.error]	Average survival in litter [std.error]
<i>Bernese Mountain Dog, L</i>	17293	535	0.8	0.63	5.7 [3.0]	5.0 [2.7]
<i>Border Collie, M</i>	16830	546	0.5	0.53	5.8 [1.3]	5.5 [2.0]
<i>Small Münsterländer, M</i>	15478	359	0.7	0.42	7.9 [2.4]	7.2 [2.4]
<i>Rottweiler, L</i>	28802	444	1.3	1.06	7.2 [3.0]	6.3 [2.8]

5.1 Distribution of Litter Size

This section will describe the distribution of litter size for the four breeds, given as the number of total born puppies for every litter.

The Bernese Mountain dog displays an unimodal distribution of litter data with a pronounced right skew. The majority of litters contain three to seven puppies, as indicated by the right skew. Litters

with more than eight puppies occur less frequently, with very few litters exceeding ten puppies. The mean litter size is 5.7 with a standard error of 3.0. The predominance of litters with a single puppy (frequency ~ 105), could potentially skew further analytical results (Appendix 3). The Border Collie displays an approximately normal distribution of litter data with a little right skew. The majority of litters contain four to five puppies (~ 130). A larger proportion of litters consist of more than five puppies, up to ten puppies, compared to those with fewer than four puppies (Appendix 4). The mean litter size is 5.8 [1.3].

The Kleiner Münsterländer displays an unimodal distribution of data with a moderate left skew. The majority of litters contain 8 puppies (~ 80) and a larger proportion of litters consist of less than 8 puppies (Appendix 5). The mean litter size is given 7.9 [2.4] puppies.

Finally, the Rottweiler displays an approximately normal distribution of data with light left skew. The majority of litters contain between 7 to 8 puppies (~ 60) with a larger proportion of litters consisting of less than 7 puppies given higher frequencies (Appendix 6). Given mean litter size 7.2 [3.0].

5.2 Pedigree Completeness

The following section will discuss the fulfillment of ancestors in the pedigrees for the four breeds. The graph in Figure 7 visualizes the development of completeness according to the number of known ancestors over time. The x-axis refers to years from 1980 to 2023 whereas the y-axis refers to the number of known ancestors given in the mean value to the exact year. The aim is to depict the trend in pedigree completeness for the studied population of the breeds in the given period of time. The mean value of pedigree completeness remained relatively stable around 1 to 2 from 1980 until the early 2000s for the most of the dogs except the Border Collie, with minor fluctuations observed during this period. However, starting in the early 2000s, there is a notable upward trend, leading to increase in pedigree completeness for all breeds. By 2023, the mean value has risen to approximately 6 for the Bernese Mountain Dog, 5 for the Border Collie, 5.5 for the Kleiner Münsterländer and the Rottweiler. For the Border Collie, the mean value of pedigree completeness remained relatively stable around 0 to 0.5 from 1980 until the 1985. A completeness of 1 is firstly reached in 1986.

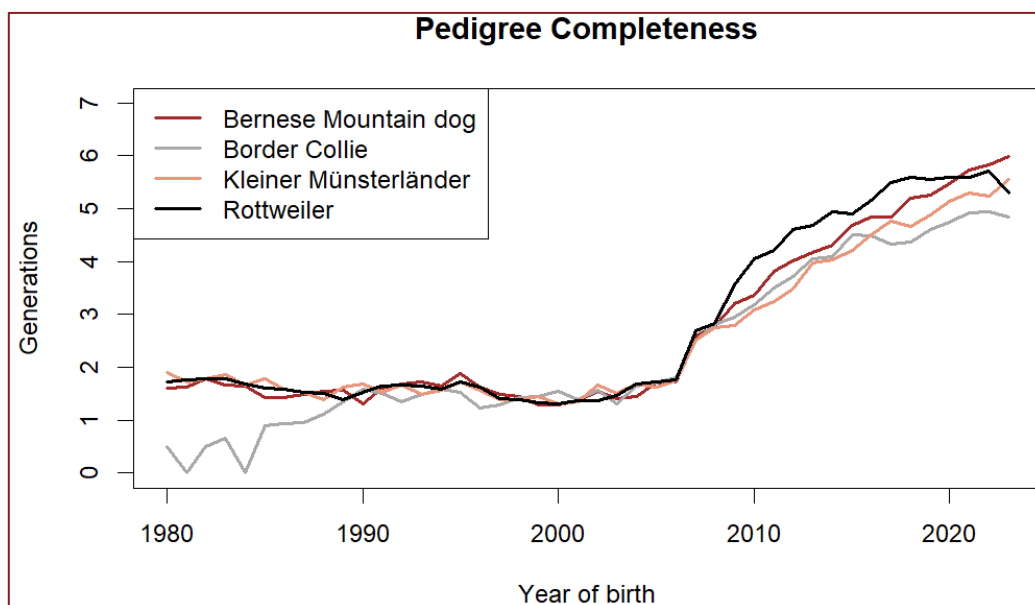


Figure 7: Line graph illustrating the trend in pedigree completeness from 1980 to 2023 for all dog breeds by the generation equivalents. The y-axis represents the generations, while the x-axis represents the year. A value of one refers to no ancestors known. Values bigger than one refers to the number of known generation ancestors.

5.3 Numbers in Pedigree

The following curve of the four breeds (Figure 8) show the number of registered dogs after birth year. The x-axis is given as the birth year ranging from 1989 to 2023. The y-axis represents the number of registered dogs ranging from zero to eight hundred.

Several fluctuations in the number of registered dogs for all the breeds is seen with distinct peaks and troughs. A significant peak for the Bernese Mountain Dog occurred around 2006, where the number of registrations reached approximately 450. Another notable peak is observed in 2020. From 2009 to 2012, a significant decline in registrations is seen, dropping from around 450 to 250. The period from 2020 to 2023 exhibits a slight decreasing trend. For the Border Collie a significant peak occurred around 2006, where the number of registrations reached approximately 450. Another notable peak is observed in 2021. From 1989 to 2006 the number increased drastically from 50 to 450 dogs. The increase from 2009 to 2023 is lower with the number close to 500 dogs in 2023. The years exhibit values between 275 and 350 dogs. The Kleiner Münsterländer shows a peak in around 1995, where the number of registrations reached approximately 470. From 1996 to 1999, there was a significant decline in registrations, dropping from around 470 to 150. From 2001 to 2023 the fluctuations stabilized around 250 to 300. The period from 2021 to 2023 exhibits a slight decreasing trend. Despite

these variations, the overall trend does not display dramatic shifts, suggesting relatively stable registration practices over the long term, however the big fluctuation in 1994 is remarkable. The most significant peak for the Rottweiler occurred around 2005, where the number of registrations reached approximately 590. Another notable peak was observed in 2014. From 1989 to 2023 the number decreased drastically from 700 to 200 dogs. The period from 2021 to 2023 exhibits a drastically decreasing trend. The overall trend displays a drastically decreasing in registration practices over the long term. The overall trend for the Bernese Mountain Dog and the Kleiner Münsterlander does not display dramatic shifts, suggesting relatively stable registration practices over the long term. Border collie visualized an increasing trend.

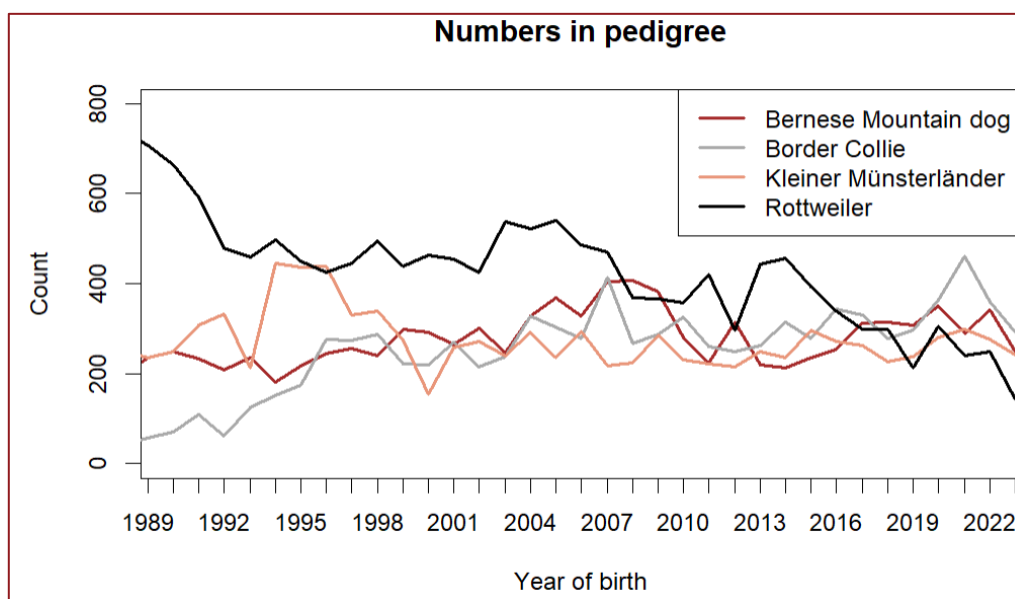


Figure 8: Shows number of dogs registered in pedigree after birth year for all the breeds from 1989 to 2023. The y-axis represents the number of puppies registered, while the x-axis represents the year of birth.

5.4 Inbreeding Trends

The genetic analysis of the four dog breeds shows generally low to moderate levels of inbreeding and genetic differentiation, along with slight excess heterozygosity within subpopulations. The completeness scores indicate good data coverage for all breeds (see Table 4). All breeds exhibit low to moderate inbreeding, with F_{it} values ranging from 0.005 in Border Collies to 0.031 in Rottweilers. Genetic differentiation among subpopulations is minimal to moderate, with F_{st} values ranging from 0.008 in Border Collies to 0.023 in Kleiner Münsterländers. Slight excess heterozygosity is observed across all breeds, indicated by negative F_{is} values ranging from -0.003 in Border Collies to -0.016 in Kleiner Münsterländers.

Table 4: Presents the genetic differentiation and inbreeding coefficients for four dog breeds, along with the completeness of the dataset used in the analysis.

Breed	Fit	Fst	Fis	Completeness
<i>Bernese Mountain Dog</i>	0.008	0.013	-0.005	5.841
<i>Border Collie</i>	0.005	0.008	-0.003	4.933
<i>Kleiner Münsterländer</i>	0.007	0.023	-0.016	5.248
<i>Rottweiler</i>	0.031	0.018	-0.004	5.718

The total number of data regarding F_{it} differ among the four breeds with a little over 100 observations in difference (Rottweiler to Border Collie) (Table 5). The mean-value represents the average value of data, which is clearly higher for the Rottweiler (0.031). Median represents the middle value in the dataset when ordered by number. That value is close among breeds, except for the Border Collie with the value equal zero, which means that for half of the animals we can't detect any inbreeding. The 90% fraction does not differ among breeds, whereas the maximum is highest for the Border Collie and Rottweiler, so those breeds detect the highest F_{it} value in 2022.

Table 5: Summarize the average values of count, mean, median, 90% fraction and maximum of F_{it} in year 2022.

Breed	Count	Mean	Median	90% frac	Max
<i>Bernese Mountain Dog</i>	342	0.008	0.002	0.02	0.01
<i>Border Collie</i>	359	0.005	0	0.02	0.07
<i>Kleiner Münsterländer</i>	276	0.007	0.004	0.02	0.03
<i>Rottweiler</i>	248	0.031	0.004	0.03	0.07

5.4.1 Wrights Coefficients

This section shows the development of Wright's F-coefficients which is depicted in the following figures 23, 24, 25 and 26, for the four breeds. F_{it} represents the total inbreeding, F_{st} as inbreeding due to drift and F_{is} as inbreeding due to non-random mating.

5.4.1.1 Bernese Mountain Dog

The inbreeding trends for the Bernese Mountain Dog (Figure 9) shows dynamic changes in genetic structure and inbreeding, illustrated as big fluctuations within the population over the years. F_{it} is close to or equal to zero from 1982 to 2009 where it increases to around 0.015 in 2020. In 1981 a little increase in F_{it} was seen. The period from 2021 to 2023 exhibits a decreasing trend in F_{it} close to 0.005. Significantly fluctuations of F_{st} and F_{is} are seen in 1994 restrictively 0.025 and -0.025. F_{st} increases slowly from 2000 to 2023 to 0.015 while F_{is} decreases to -0.005 in 2019. Both F_{it} and F_{is} decrease from 2020 to 2023.

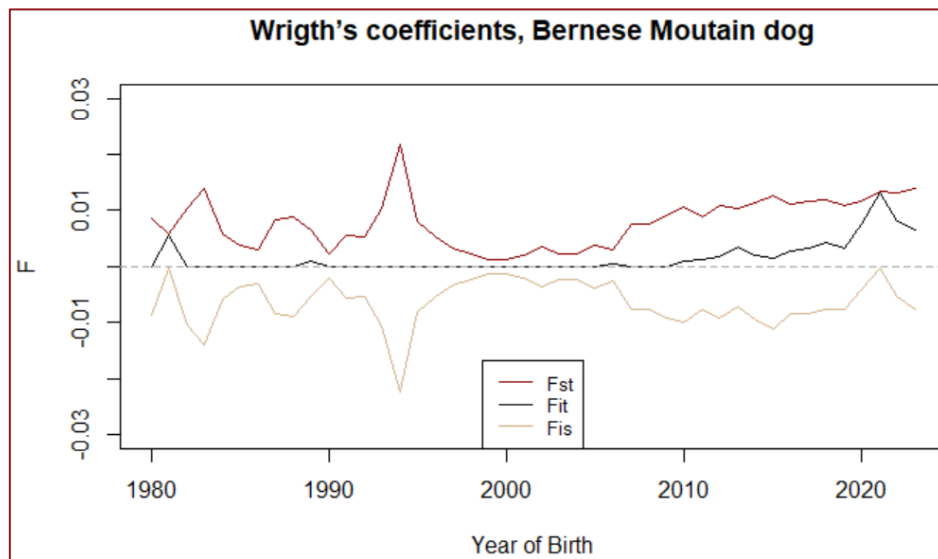


Figure 9: Trends in Wright's F-statistics (F_{st} , F_{it} , and F_{is}) from 1980 to 2023 for the Bernese Mountain Dog. The y-axis represents the F values, while the x-axis represents the year of birth. The red line shows the fixation index (F_{st}), indicating genetic differentiation among subpopulations. The black line represents the inbreeding coefficient relative to the total population (F_{it}), and the tan line represents the inbreeding coefficient relative to the subpopulation (F_{is}).

5.4.1.2 Border Collie

The inbreeding trends for the Border Collie (Figure 10) shows changes in genetic structure and inbreeding, illustrated as less fluctuations within the population over the years in comparison to the Bernese Mountain Dog. F_{it} is close to or equal to zero from 1986 to 2007 where it increases to around 0.01 in 2017 with two fluctuations in 2011 and 2017 followed with a decrease in 2018. The period from 2021 to 2023 exhibits an increasing trend in F_{it} close to 0.01. F_{is} follows the same pattern as F_{it} from 2007 followed by increased fluctuations in 2011 and 2017. F_{st} and F_{is} follows the same mirror pattern from 1986 to 2008. Big fluctuations of both are visible in 1990 and 1992 with respectively 0.015 and -0.012. All coefficients increase slowly from 2000 to 2023.

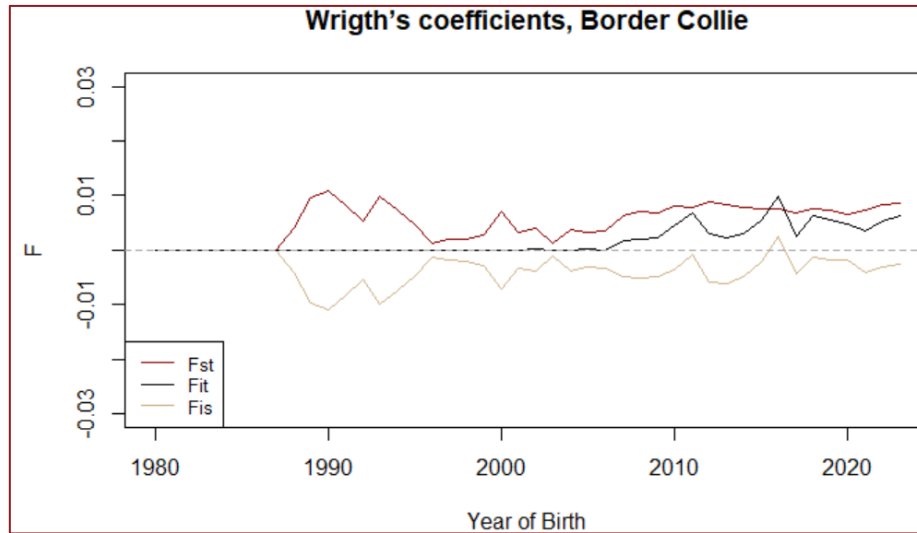


Figure 10: Trends in Wright's F-statistics (F_{st} , F_{it} , and F_{is}) from 1980 to 2023 for the Border Collie. The y-axis represents the F values, while the x-axis represents the year of birth. The red line shows the fixation index (F_{st}), indicating genetic differentiation among subpopulations. The black line represents the inbreeding coefficient relative to the total population (F_{it}), and the tan line represents the inbreeding coefficient relative to the subpopulation (F_{is}).

5.4.1.3 Kleiner Münsterländer

The inbreeding trends for the Kleiner Münsterländer (Figure 11) shows changes in genetic structure and inbreeding, illustrated as later fluctuations within the population over the years. F_{it} is close to or equal to zero from 1980 to 2007 where it increases to around 0.006 in 2022 with two little fluctuations in 2008 and 2015. The period from 2021 to 2023 exhibits a stabilized trend in F_{it} close to 0.005. F_{st} and F_{is} follows the same mirror pattern from 1983 to 2023. Big fluctuations of both are visible from 2005 and forward with respectively 0.03 and -0.025. F_{is} continues to decrease from 2021 to 2023 while F_{st} and F_{it} increases and stabilizes.

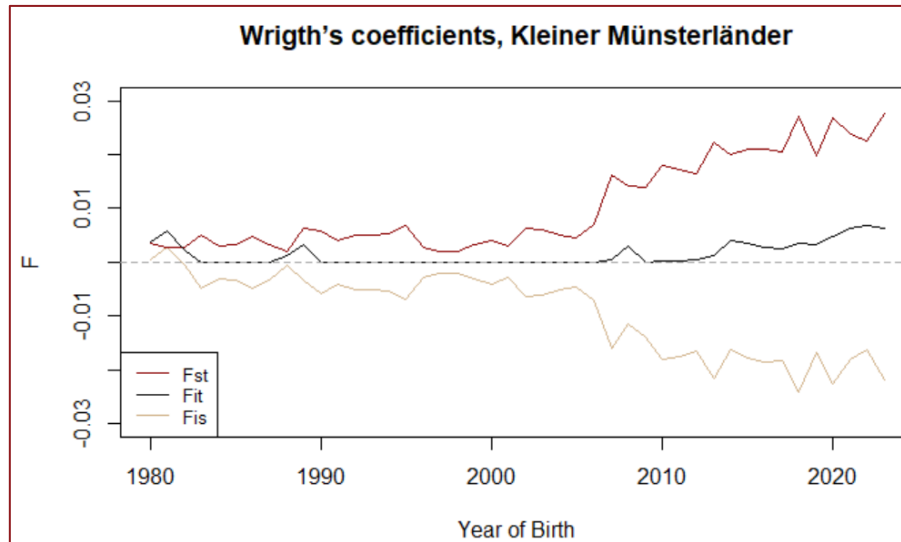


Figure 11: Trends in Wright's F-statistics (F_{st} , F_{it} , and F_{is}) from 1980 to 2023 for the Kleiner Münsterländer. The y-axis represents the F values, while the x-axis represents the year of birth. The red line shows the fixation index (F_{st}), indicating genetic differentiation among subpopulations. The black line represents the inbreeding coefficient relative to the total population (F_{it}), and the tan line represents the inbreeding coefficient relative to the sub-population (F_{is}).

5.4.1.4 Rottweiler

The inbreeding trends for the Rottweiler (Figure 12) shows changes in genetic structure and inbreeding, illustrated as later fluctuations within the population over the years. F_{it} is close to or equal to zero from 1980 to 2009 where it increases to around 0.018 in 2021 with fluctuations in 2009 and 2013. The period from 2021 to 2023 exhibits a decreasing trend in F_{it} close to 0.004. F_{it} and F_{st} increases from 2009 to 2021 where F_{st} keeps increasing while F_{it} exhibits a decreasing trend. Fluctuations are seen almost all years from 2006 to 2023 for all coefficients. F_{is} decreases slowly from 2006 to 2023, from close to zero to -0.01, but remains around stable in that period between zero and -0.01.

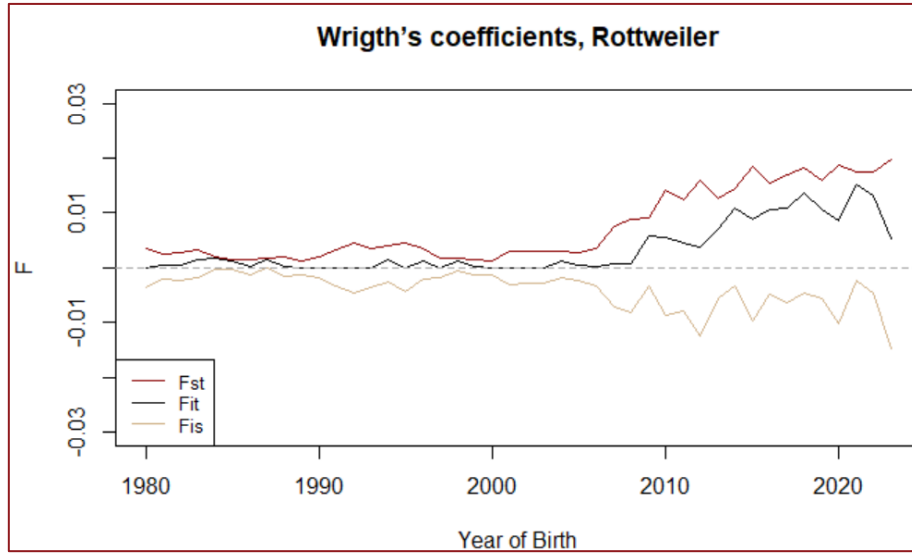


Figure 12: Trends in Wright's F-statistics (Fst, Fit, and Fis) from 1980 to 2023 for the Rottweiler. The y-axis represents the F values, while the x-axis represents the year of birth. The red line shows the fixation index (Fst), indicating genetic differentiation among subpopulations. The black line represents the inbreeding coefficient relative to the total population (Fit), and the tan line represents the inbreeding coefficient relative to the subpopulation (Fis).

5.5 Statistical Analysis

The aim of the following analysis is to investigate the correlation between litter size as the dependent variable and the explanatory variables, which include year, age of the mother, inbreeding coefficients of the mother, father, and offspring, and parity. A linear regression analysis was conducted in R (lm) to analyze the data.

For all four breeds, a linear regression was performed including all the factors (full_model):

$$\text{Litter size} = \beta_0 + \beta_1 * \text{year} + \beta_2 * \text{dam.age} + \beta_3 * \text{daminb} + \beta_4 * \text{offinb} + \beta_5 * \text{sireinb} + \text{parity} + \epsilon$$

Litter size is the dependent variable, while the others are explanatory variables. ϵ represents the residuals, accounting for the variation in litter size that cannot be explained by the explanatory variables. The results can be seen in Table 6. The model will be simplified by removing factors that have no significant effect to investigate the following null hypothesis:

H0: Year has no effect on the litter size of the four dog breeds.

H0: Inbreeding of sire has no effect on the litter size of the four dog breeds.

H0: Inbreeding of offspring has no effect on the litter size of the four dog breeds.

H0: Age of dam has no effect on the litter size of the four dog breeds.

H0: Parity has no effect on the litter size of the four dog breeds.

Year has no significance effect on any of the models. Age of dam is highly significant for all models ($p < 0.05$). The inbreeding of dam is only significant for the Rottweiler ($p < 0.05$). Inbreeding of offspring shows no significant effect on litter size either ($p > 0.05$) in any of the models. The inbreeding of sire has no effect on any of the models ($p > 0.05$). The parity number has a significant effect on litter size only for the medium sized breeds ($p < 0.05$).

The variance analysis of the models shows no statistical significance differences between any of the two models analyzed, but a possible difference between M1_km and M2_km for the Kleiner Münsterländer by anova ($p=0.05$). The variable “year” is still removed due to no-significant effect.

According to the estimates, the age of the dam has negative effects due to negative values for all the models (estimates < 0), while parity has a positive effect (estimates > 0) except for the Bernese Mountain Dog (estimates < 0). Regarding the Rottweiler, the inbreeding of dam effects litter size negatively (estimates < 0).

The simplified models are given:

$$Litter\ size_{Bernese\ Mountain\ Dog} = \beta_0 + \beta_1 * dam.age + \epsilon$$

$$Litter\ size_{Border\ Collie, Kleiner\ Münsterländer} = \beta_0 + \beta_1 * dam.age + \beta_2 * parity + \epsilon$$

$$Litter\ size_{Rottweiler} = \beta_0 + \beta_1 * dam.age + \beta_2 * daminb + \epsilon$$

Table 6: Shows the investigated models from full model to simplified and all the explanatory variables estimates / p-values for the analysis of all four breeds. The p-value of the variance analysis (anova) is seen to the right, given in p-values for the model and the model before. The crosses refers to the removed variables in the model. The stars refers to the significance by the p-value < 0.05 .

Model/explanatory variable (test)	Year	Dam.age	Daminb	offinb	Sireinb	Parity	(anova)
Bernese Mountain Dog (M1_bs)	3.34 * 10 ⁻² / 0.47	-1.43 * 10 ⁻³ / 1.19 * 10 ⁻⁵ ***	-1.84 / 0.11	-2.24/ 0.79	8.41/ 0.94	-6.74 * 10 ⁻² / 0.73	
M2_bs	x	-1.48 * 10 ⁻³ / 3.67 * 10 ⁻⁶ ***	-1.76/ 0.12	-1.59 / 0.85	0.95 / 0.94	-3.33 * 10 ⁻² / 0.86	0.47
M3_bs	x	-1.48 * 10 ⁻³ / 3.6 * 10 ⁻⁶ ***	-1.75 / 0.12	-1.55 / 0.85	x	-3.32 * 10 ⁻² / 0.86	0.94

M4_bs	x	$-1.47 * 10^{-3}$ / $3.59 * 10^{-6}$ ***	$-1.79 /$ 0.11	x	x	$-3.42 *$ $10^{-2} / 0.86$	0.85
M5_bs	x	$-1.51 * 10^{-3}$ / $9.66 * 10^{-9}$ ***	$-1.78 /$ 0.11	x	x	x	0.86
M6_bs	x	$-1.49 * 10^{-3}$ / $1.43 * 10^{-8}$ ***	x	x	x	x	0.11
Border Collie (M1_bc)	$3.28 *$ 10^{-2} /0.27	$-4.02 * 10^{-4}$ / 0.01*	$-5.79 / 0.36$	$3.85 / 0.42$	$1.89 / 0.699$	$0.25 /$ $0.0200 *$	
M2_bc	x	$-4.28 * 10^{-4}$ / 0.008 **	$-6.26 /$ 0.32	3.78 / 0.43	$1.81 / 0.71$	$0.28 / 0.007$ **	0.27
M3_bc	x	$-4.31 * 10^{-4}$ / 0.007 **	$-6.31 /$ 0.31	3.73 / 0.43	x	$0.29 / 0.006$ **	0.94
M4_bc	x	$-4.28 *$ $10^{-4} /$ 0.008**	-5.64 / 0.36	x	x	$0.29 /$ $0.006 **$	0.43
M5_bc	x	$-4.15 * 10^{-4}$ / 0.01*	x	x	x	$0.28 / 0.007$ **	0.36
Kleiner Münsterländer (M1_km)	$-8.83 *$ $10^2 /$ 0.05.	$-1.28 *$ $10^{-3} / 5.34 *$ $10^{-6} ***$	$14.2 / 0.4$	$22.5 / 0.17$	-0.26 / 0.99	$0.87 /$ $3.34 * 10^{-5}$ ***	
M2_km	x	$-0.001 /$ $1.67e-05 ***$	$9.36 / 0.58$	$17.90 /$ 0.28	$-2.35 / 0.9$	$0.77 /$ $0.0001 ***$	0.05 .
M3_km	x	$-0.0004 /$ 0.007 **	$-6.31 / 0.31$	$3.73 / 0.44$	x	$0.29 / 0.007$ **	0.91
M4_km	x	$-0.001 / 1.05 *$ $10^{-5} ***$	$10.75 / 0.52$	x	x	$0.77 /$ $0.0001 ***$	0.28
M5_km	x	$-0.001 / 4.4 *$ $10^{-6} ***$	x	x	x	$0.79 / 7.75 *$ $10^{-5} ***$	0.52
Rottweiler (M1_r)	$3.04 *$ $10^{-2} /$ 0.56	$-1.03 *$ $10^{-3} / 0.003 **$	$-19.49 /$ 0.01 *	$4.92 / 0.52$	$-3.3 / 0.75$	$0.19 / 0.26$	
M2_r	x	$-1.04 * 10^{-3}$ / 0.002**	$-19.41 /$ 0.01 *	$4.95 / 0.51$	$-2.38 / 0.82$	$0.21 / 0.20$	0.56

M3_r	x	-1.04 * 10 ⁻³ / 0.00224 **	-19.33 / 0.012 *	4.73/ 0.53	x	0.21 / 0.21	0.81
M4_r	x	-1.05 * 10 ⁻³ / 0.002**	-18.89/ 0.02*	x	x	0.21 / 0.21	0.53
M5_r	x	-8.09 * 10 ⁻⁴ / 0.003**	-18.18 / 0.02 *	x	x	x	0.21

The effect of maternal age is greatest for the Bernese Mountain Dog with 0.54 fewer pups in litters per year, and lowest for the Border Collie with 0.15 fewer pups (Table 7).

Table 7: Showing the effect of age of dam on litter size for a year, indicating the lower numbers of puppies per litter due to the effect of dams age.

Breed	Dam.age
Bernese Mountain Dog	-0.54
Border Collie	-0.15
Kleiner Münsterlän- der	-0.37
Rottweiler	-0.30

6. Discussion

The following section discusses the findings of the analysis in the context of the results with the existing literature, on the background of the theoretical frame. Due to the limited research on the effects of inbreeding in relation to breed and litter size, studies that met specific criteria were included to narrow the scope. The purpose of this study was to understand the relationship between inbreeding, body size and litter size.

This research has introduced new perspectives on the effects related to the influence of litter size, and the differential effects among two body sizes of four different breeds. The results of the statistical analysis indicate that the factors affecting litter size vary between breeds, which is consistent with previous studies (McMillian et al., 2024; Kraus et al., 2023; Kania-Gierdziewicz & Pałka, 2019).

6.1 Effects on Litter Size

The explanatory variables investigated were analyzed by the null hypothesis (H0), which will be discussed.

H0: *Age of dam has no effect on the litter size of the four dog breeds.*

According to Borge et al. (2011), litter size is influenced by more than one factor (age of dam, size of breed and method of mating) whereas the age of dam significantly influenced litter size ($p < 0.01$) for all the breeds of varying sizes across the 224 breeds studied. The results of this study agree with the age of dam as an explanatory variable describing litter size in the four breeds studied ($p < 0.05$). However, Schelling et al. (2019) disagree, as the analysis of variance conducted in the study did not show a significant effect of dam age on litter size. Sample size and diversity could be the reason for the different results, as Borge et al. (2011) used 224 different breeds, Schelling et al. (2019) only studied Labrador Retrievers, while this study only focused on four different breeds. A larger and more diverse sample size would provide more robust statistical power and generalizability, moreover different breeds or groups of breeds may have unique reproductive characteristics being more specific to one breed. Therefore the null hypothesis (H0) regarding the impact of the age of dam on litter size can be discarded, due to the broader results and larger samples for Borge et al. (2011) and this study.

H0: *Parity has no effect on the litter size of the four dog breeds.*

Borge et al. (2011), figured by an analysis on unconditional correlations that litter size decreased with increasing parity ($p < 0.01$). This could possibly be due to a direct effect of parity. Even so, the

following multiple variable analysis showed that the tendency primarily was due to the increasing age of the dam since the effect got insignificant by involving the age of the dam. That was not the case in this study. The results of this study indicated that parity only affects medium sized dogs. A decrease in litter size was observed with increased parity. Also, the multivariable model of this study included both factors to be highly significant for the two medium sized breeds. Borge et al. (2011) suggest that the negative effect of parity on litter size is confounded by dam age. Differences in results could be due to the possible different relationship between parity and litter size, which can vary significantly between breeds. The discrepancy in the sample size of the investigated breeds could potentially contribute to the observed variations in the results. Borge et al. (2011) included a larger and more diverse sample size, which may have led to different conclusions than this more focused study on medium-sized and large breeds, which included only two breeds in each category. Therefore the null hypothesis (H0) regarding the impact of the age of dam on litter size can be partly discarded because it does not affect all breeds and body sizes, but only medium sized dogs according to the results in this study and could possibly be confounded to the age of dam according to earlier studies.

H0: *Inbreeding of sire has no effect on the litter size of the four dog breeds.*

Kania-Gierdziewicz & Pałka, (2019) observed that the inbred sire resulted in a higher proportion of female puppies, which was attributed to the deterioration in the sex ratio. This suggests a possible impact of sire-specific traits on litter traits. In contrast, other studies have indicated that it is the genetic traits of the dam that affect litter traits, particularly her productive capabilities, which in turn affect litter size (Leroy et al., n.d). Nevertheless, Leroy et al. (n.d.) determined that the inbreeding coefficient of the sire had a significant effect on the litter size of the two breeds investigated ($p = 0.04$, $p = 0.007$), although these findings cannot be directly compared to the four breeds included in this study. Schelling et al. (2019) demonstrated that the sire had no effect on the litter trait “body weight at birth.” Moreover, Marelli et al. (2020) demonstrated that an increase of 1% inbreeding in the sire resulted in a 0.007% increase in the number of dead puppies in the litter, a comparatively minor effect compared to 0.02% increase in the same inbreeding of the dam. The statistical analysis of the pedigrees revealed that there was no evidence to suggest that the inbreeding level of the sire had an impact on litter size ($p > 0.05$). The observed discrepancy may be a result of the specific breeds used in this study, the sample size, or other uncontrolled environmental factors that differ from those in other studies regarding data. Therefore the null hypothesis (H0) regarding the impact of the inbreeding level on litter size cannot be discarded.

H0: *Inbreeding of offspring has no effect on the litter size of the four dog breeds.*

The impact of inbreeding on offspring's effect on litter size has been the subject of several investigations in studies. Marelli et al. (2020) observed a slight increase in the number of dead puppies born when the offspring was inbred with 1% in comparison with the same inbreeding of the dam. This resulted in a 0.06% increase in the number of dead puppies. A number of studies have indicated that inbreeding of puppies has a negative effect on litter size (de Andrade et al., 2021; Leroy et al. n.d.). de Andrade et al. (2021) have expressed that the correlation does not occur in a linear fashion as it has been described and used in the literature. Even so, Kania-Gierdziewicz & Palka (2019) found the correlation coefficients between litter inbreeding and litter size for the majority of examined dog breeds were positive but not significant ($p > 0.05$). The results in this study showed no correlation between inbreeding in offspring and litter size. Differences could be due to the understanding that the variables occur linearly, why all the studies investigated litter size or inbreeding that way. de Andrade et al. (2021) claims that the correlation between the inbreeding of offspring on litter size is not necessarily linearly, why different statistical methods could lead to different conclusions. The Study by de Andrade et al. (2021) included a small sample size of 105 litters from a single breed (German Spitz dogs), with data from a single year. The results are not particularly reflective of a larger population. Kania-Gierdziewicz & Palka's 2019 analysis of five dog breeds and 436 litters offers a more representative data analysis and results. Leroy et al. (n.d.) selected seven breeds to cover a broad range of morphological characteristics, utilization and demographic patterns with a substantial sample size and comprehensive pedigrees. Therefore the H0 cannot be discarded with the results in this study, but according to the literature, it somehow could for some breeds. Essentially, you need to take into account methods and sample sizes as well as statistical methods to compare across included studies.

H0: *Year has no effect on the litter size of the four dog breeds.*

According to the literature analyzed in this study, are there no indications that year would affect litter size since it is not investigated as an explanatory variable on the litter size as this study did it. However, the year of birth is known to potentially affect birth weight due to factors such as climate changes, management practices, or the selection of domestic species (Schelling et al., 2019). This raises the possibility that the year of birth could also influence litter size, particularly if breeding criteria have been modified to prioritize the health of the animals or other significant factors. By the statistical method in this result section it is clear that year has no significant effect on litter sizes for

the four investigated dog breeds ($p > 0.05$), therefore the null hypothesis (H_0) regarding the impact of the year on litter size cannot be discarded.

Breed and Body Size

The litter sizes of two body sizes were found to be influenced by distinct factors in this study. Both body sizes were influenced by the age of the dam, with the medium-sized dogs additionally affected by parity. Borge et al. (2011) revealed a curvilinear relationship between age and litter size, with a significant interaction between age and breed size. This indicates that the effect of maternal age on litter size varies across different sizes of breeds. Specifically, in small dogs, both young and old bitches produced smaller litters compared to those of intermediate age. In contrast, for larger breeds, a low maternal age did not reduce mean litter size, but the number of puppies per litter steadily decreased with increasing age. The mean litter size increases with breed size in Borge et al. (2011), which will be discussed in the next part, but according to figure five and six in this study, the results indicate a correlation between an increase in body size and an increase in litter size. Leroy et al. (n.d.) concluded that inbreeding depression being more pronounced in larger breeds and that can be linked to their larger litter sizes. Indicating that the level of inbreeding is significantly related to body size, with inbreeding levels increasing as body size increases ($r = 0.18$, $p = 0.02$). The research by Yordy et al. (2020) corroborates this, showing a stronger correlation between body size and inbreeding in the female line (mtDNA) compared to the male line (chrY). The data in this study do not align with the results obtained in previous studies, as the level of inbreeding among the four breeds was not as high as any of the breeds previously studied. Furthermore, only the Rottweiler exhibited signs of inbreeding due to the dam with the highest inbreeding coefficient among the four studied breeds. The findings of this study are consistent with those of Yordy et al. (2020), indicating that the impact of dams is positive, while the effect of sire is negative.

6.2 Inbreeding Trends

According to the investigated data from DKK, the average inbreeding coefficient did not vary much among the four breeds, except for the Rottweiler with a bit higher value ($F = 1.3$) (Table 3). The average inbreeding coefficient for the litter of the Bernese Mountain Dog was estimated to be 0.63%, indicating a low inbreeding value. This result differs from previous studies, in which Leroy et al. (n.d.) calculated an inbreeding coefficient of 2.08%. The difference could be due to the fact they considered litters from 1990 to 2012 with at least three equivalent generations of known ancestors, the results in this study rely on ten years of registered litters ranging from 2013 to 2023. The countries

are different with different genetic, breeding practices and environment which could impact the possibility of correlated results. The rules for breeding can also vary although some main rules could be found.

Since the included four dogs did not tend to have a high inbreeding coefficient (between 0.8 to 1.3), it was not possible to conduct a significant influence of so on litter size. According to the results, only the Rottweiler seemed to have an effect on the litter size by the inbreeding of the dam (-18.18, $p > 0.05$). Studies do indicate that high inbreeding is correlated with a decrease in litter size (de Andrade et al., 2021; Leroy et al., n.d.), but those results from the analysis in this project cannot clarify that claim directly with low inbreeding coefficients. Nevertheless, the findings demonstrate that the inbreeding of the mother has a negative effect on litter size in the Rottweiler with the highest but not significantly high inbreeding coefficient. Since more studies agree to the inbreeding of dams to affect litter size (Marelli et al., 2020; Kania-Gierdziewicz & Pałka, 2019), these low inbreeding coefficients could be too low to conduct a possible correlation. Kania-Gierdziewicz & Pałka (2019) did find a positive correlation between those factors, but it was not significant ($p > 0.05$). No indication of a correlation was to be seen in this study.

The inbreeding tendencies will now be discussed to the three investigated categories, F_{st} , F_{it} and F_{is} in correlation with pedigree completeness and numbers of registered:

The number registered in the pedigree is fairly stable at around 200 to 300 dogs each year for the Bernese Mountain Dog and the Kleiner Münsterländer. The Rottweiler has been in steep decline since 1989, while the Border Collie has increased in registrations over the years. The interest in the Bernese Mountain Dog and Kleiner Münsterländer has remained relatively unchanged, while it has decreased for Rottweilers and increased for Border Collie (Figure 8). In 2010 a new law came into force in Denmark prohibiting 13 so-called “dangerous” dog breeds (Ministeriet for Fødevarer, Landbrug og Fiskeri, 2021). Rottweiler was not among them, but the breed had appeared on a gross list prior to the passing of the law. This perception in public of the Rottweiler as a potentially dangerous breed might have influenced the popularity of the breed. The Border Collie on the other hand has been suggested to be an intelligent dog breed in number of ratings (cognitive performance, point-following tasks, quantity and discrimination tasks) which may have had an impact on the popularity of the breed (Arden & Adam, 2016). From 2005 to 2023, the generation has increased from 1.5 to 5.5. Even so the completeness of pedigree has increased, the number of foreign dogs affects the inbreeding coef-

ficients and thereby Wrights coefficients by overshadowing the values due to less pedigree information.

An increase in pedigree completeness signifies improvements in the accuracy, depth, and comprehensiveness of the genealogical records in the pedigree. Newer records include more generations which allows for deeper understanding of the genetic backgrounds of the dogs within their population. When investigating the F_{is} , F_{st} and F_{it} , the mean values of the relevant year were taken into account, so the first years rely on a low pedigree completeness.

6.3 Distribution of Litter Size

The distribution of litter sizes varies among the four breeds. The Bernese Mountain Dog contributes a high frequency of litter sizes with one puppy (figure 7). This occurrence is typically uncommon and may suggest an issue, either with the registration process or with the breed itself. It is well-known that litters of one puppy complicates the birth for the mother, since the offspring is not able to secrete oxytocin which starts birth that is why those incidents would demand cesarean section. For the Bernese Mountain Dog, the incidence of big litter sizes is rare but seen. The average litter size was calculated to be 5.7 [± 3.0] (table 3), whereas elsewhere it has been calculated as 5.51 [± 2.78] (Leroy et al., n.d.) and 6.4 [± 0.3] (Borge et al., 2011). The Border Collie shows a more expected distribution of litter sizes with most litters containing four to six puppies. Interestingly the litters could be higher for that breed up to seven but also giving birth to litters with one puppy (figure 8) The average litter size was calculated to be 5.8 [± 1.3] (table 3), differently from Borge et al. (2011), whereas it was 6.0 [± 0.1]. The Kleiner Münsterländer typically gives birth to litters averaging around eight puppies. While larger litters are occasionally observed, most litters consist of fewer than eight puppies < 8 (figure 9). Borge et al. (2011), calculated the average litter size to 6.9 [± 0.5] whereas this study got 7.9 [± 2.4] (table 3). The Rottweiler gave birth to litters that are more evenly distributed from one to ten puppies, but a bit higher tendency is seen with smaller litter sizes (figure 10) and given the mean value at 7.2 [± 3.0] (table 3), which is consistent with the findings Borge et al. (2011), who reported a mean value of 7.4 [± 0.2].

Very few of the studies calculate the inbreeding coefficient, which makes it challenging to discuss the results from the research in that perspective. Borge et al. (2011), calculated the same value for the same breeds with a coefficient that was almost identical, except for the Kleiner Münsterländer, which exhibited a slightly higher value. The coefficient varies among breeds and is affected by the pedigree completeness.

Borge et al. (2011), corroborates the findings of this study by demonstrating that mean litter size increases with breed size. This relationship was statistically significant in a simple linear regression model, showing mean litter sizes of 3.5 in miniature breeds, 4.2 in small breeds, 5.7 in medium breeds, 6.9 in large breeds, and 7.1 in giant breeds. This gradation underscores the importance of breed size in reproductive outcomes, aligning closely with our results. This is not consistent with our own results, as the Kleiner Münsterländer, a medium-sized dog, had a significantly larger average litter size compared to the Bernese Mountain Dog with large body size. However, it is important to consider the high frequency of litters with 1 puppy for the Bernese Mountain Dog.

6.4 Methods

The purpose of this section is to evaluate the methods used in the current studies and to assess their strengths and weaknesses.

The methods used in the literature are for the most part consistent with the methods used in this study. The majority of studies employ a pedigree-based approach to statistical analysis, utilizing data from a diverse range of countries and years. This methodology is advantageous in the context of complex data and the identification of patterns in large datasets. Geographical variation can lead to significant variation in data sources and quality, making it challenging to make meaningful comparisons. Data collected at different times can be used to identify changes over time, such as seasonal variations or temporal trends. The pedigree information ranges from national kennel clubs to publicly available databases. Litter varies among studies from 10.810 (Borge et al., 2011), to 1.013 (Schelling et al., 2019) while the data in this study included 1.984 litters in total. The number of dog breeds can affect the credibility and conclusions of how broadly you can draw conclusions from the results, and those vary among studies from 9 different breeds including the Border Collie (McMillian et al., 2024) to 7 breeds (Leroy et al., n.d) to 117 breeds (Yordy et al., 2020) and to 224 breeds (Borge et al., 2011). This study only included four different breeds in two different body classes.

A strength of the methods used is their precision in measuring relationships between variables. For example, regression analysis allows precise quantification of the relationship between maternal age and litter size. The methods are also relevant to the research questions as they can test hypothetical relationships in a large population as done in this study, and the same methods create scientific credibility and make them comparable.

A common limitation of the methods is the potential bias in data collection, such as selection bias, which can affect the results. In addition, the use of self-reported data in pedigree analysis can introduce errors. Furthermore, the statistical models may have limitations in handling all possible confounding factors. When conducting statistical correlations, the choice of parameters and assumptions on data can affect the outcome of the analysis.

Compared to alternative methods, such as experimental designs, the methods used provide broader generalizability, but at the cost of control over variables. Experimental methods could have provided more controlled results, but would be less practical in a large population study.

6.5 Challenges in This Study

6.5.1 Scientific Uncertainty

One of the primary limitations of this study is the empirical uncertainty due to the scope and length constraints of the assignment, since it limits the number of breeds included. Therefore, the overall reliability of the findings would be enhanced by including larger sample sizes of dogs, including more breeds with different body sizes. The current data set, comprising a limited number of individuals in each group, introduces a degree of empirical uncertainty, whereas an expanding to more breeds would provide a more comprehensive and reliable picture, reducing the impact of sampling errors and increasing the robustness of the results. The theoretical uncertainties implies a restricted depth of exploration into other potentially relevant factors. For instance, while mating behaviors were included in the dataset, a thorough analysis was not feasible due to the lack of comprehensive theoretical knowledge and the scope of the study. Including a broader theoretical framework and more extensive literature review, including meta-studies, could provide deeper insights into these additional factors and their potential impact on the results.

6.5.2 Confounding and Causality

Possible presence of confounding variables, such as seasonal effects and environmental conditions, possess another challenge. Each breed's unique characteristics require a detailed understanding to accurately assess these confounders. The study's design also raises questions about causality. Specifically, it is crucial to ascertain whether the study genuinely investigates the intended questions and factors. By statistical analysis on the basis of the dataset, the investigated variables are quite clear which suggest that the analysis is consistent with the research question, even so some factors are beyond that. The findings in this study indicate that body size is hypothesized to be linked to larger

litter sizes (Figure 5 and 6), in agreement with Borge et al. (2011). This could be due to the relative size of the fetus, which is bigger in larger-sized breeds, and thereby a confounder (Borge et al., 2011). Moreover age is associated with smaller litter sizes. Ensuring that the study's design and data collection methods effectively capture these causal relationships is vital for the validity of the findings. Additional studies focusing on these causal links, supported by robust theoretical frameworks, are necessary to confirm these hypotheses. To mitigate the impact of confounding variables, the study attempted to control as many variables as possible by simplifying the statistical models. However, it is acknowledged that not all variables could be accounted for, and some residual confounding may remain.

The findings presented here represent a step in building a broader understanding but should be corroborated by higher levels of evidence, such as systematic reviews and meta-analyses, to provide more definitive conclusions. Integrating the study's results with existing literature through evidence synthesis methods will help establish a more robust understanding of the factors influencing litter size in dogs.

7. Conclusion

Based on the analysis, the results and discussion, it can be concluded that there is a significant correlation between dam age and litter size. The answer to the research question is that the inbreeding level of the four dogs was probably not high enough to conclude an effect, and the inbreeding trends were not revolutionary. For all breeds, dam age had a negative effect on litter size. In addition, the two different body sizes had different effects on litter size, while the medium size dogs (Border Collies and Kleiner Münsterländer) were also affected by parity. Some studies, included in this research, launched the possible confounding effect between dam age and parity, which should be taken into account. The Rottweiler was the only breed whose litter size was affected by the inbreeding of the dam. This could be a result of a slightly higher inbreeding coefficient, which may explain why no similar findings could be made for other breeds. These findings have significant implications for the breeding of older dams, and therefore, the potential negative impact on litter size, which may result in reduced earnings. While this research provides new insights into the impact of litter size in dogs, it is important to consider the potential complications that may have affected the results. These include the lower sample size, the number of breeds included, and the pedigree completeness that is affected by foreign dogs.

This also applies to the absence of in-depth investigation of certain fact patterns, including mating, as well as the low knowledge of environmental influences and implications.

Literature

- Arden, R., & Adam, M. J. (2016). A general intelligence factor in dogs. In *Intelligence*, 50, 79-85. Elsevier. <https://doi.org/10.1016/j.intell.2016.01.008>
- Barton, N. H., & Keightley, P. D. (2002). Understanding quantitative genetic variation. In *Nature Reviews Genetics* (Vol. 3, Issue 1, pp. 11–21). <https://doi.org/10.1038/nrg700>
- Berner Sennen Klubben. (n.d.). Hvalpeliste. Retrieved June 6, 2024, from <https://berner-sennen.dk/kob-af-hvalp/hvalpeliste/>
- Borge, K. S., Tønnessen, R., Nødtvedt, A., & Indrebø, A. (2011). Litter size at birth in purebred dogs- A retrospective study of 224 breeds. *Theriogenology*, 75(5), 911–919. <https://doi.org/10.1016/j.theriogenology.2010.10.034>
- Calboli FC, Sampson J, Fretwell N, Balding DJ. (2008). Population structure and inbreeding from pedigree analysis of purebred dogs. In *Genetics*. May;179(1):593-601. doi: 10.1534/genetics.107.084954.
- Dansk Kennel Klub. (n.d.-a). *Etiske anbefalinger for avl*. Dansk Kennel Klub. Retrieved April 23, 2024, from <https://www.dkk.dk/alle-emner/opdr%C3%A6t/etiske-anbefalinger-til-opdr%C3%A6t>
- Dansk Kennel Klub. (n.d. -b). *Hundeleksikon*. Dansk Kennel Klub. Retrieved May 10, 2024, from <https://www.dkk.dk/race>
- Dansk Kennel Klub. (n.d.-c). *Om foreningen*. Dansk Kennel Klub. Retrieved April 23, 2024, from <https://www.dkk.dk/om-dkk/om-foreningen>
- de Andrade, F. M., da Silva, M. M., Krebs, G., Feltes, G. L., & Cobuci, J. A. (2021). Inbreeding on litter size of German Spitz dogs. *Revista Brasileira de Zootecnia*, 50, 1–9. <https://doi.org/10.37496/rbz5020200083>
- Ekstrøm, C.H., Sørensen, H (2015). *Statistical Data Analysis for the Life Sciences* (second edition). CRC Press – Taylor & Francis group.
- Ministeriet for Fødevarer, Landbrug og Fiskeri. (2021). *Bekendtgørelse af lov om hunde* [LBK nr. 329]. Fødevarestyrelsen. Retrieved May 20, 2024, from <https://www.retsinformation.dk/eli/lt/2021/329>

- Galibert, F., Quignon, P., Hitte, C., & André, C. (2011). Toward understanding dog evolutionary and domestication history. In *Comptes Rendus - Biologies* (Vol. 334, Issue 3, pp. 190–196). Elsevier Masson SAS. <https://doi.org/10.1016/j.crvi.2010.12.01>
- Guldbrandtsen, B. (2023). *Kompendium i Almen Husdyravl*. Institut for Veterinær- og Husdyrvidenskab.
- Hartl, D. L. (2020a). Genetic Polymorphisms. In *A Primer of Population Genetics and Genomics* (pp. 1–20). Oxford University Press. Oxford. <https://doi.org/10.1093/oso/9780198862291.003.0001>
- Hartl, D. L. (2020b). Inbreeding and Population Structure. In *A Primer of Population Genetics and Genomics* (pp. 47–74). Oxford University Press. Oxford.
<https://doi.org/10.1093/oso/9780198862291.003.0003>
- Hartl, D. L. (2020c). Population Genetics of Complex Traits. In *A Primer of Population Genetics and Genomics* (pp. 225–262). Oxford University Press. Oxford.
<https://doi.org/10.1093/oso/9780198862291.003.0008>
- Hill, W. G. (2017). Selective Breeding. In *Reference Module in Life Sciences*. Elsevier.
<https://doi.org/10.1016/b978-0-12-809633-8.07124-7>
- Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: Defining, estimating and interpreting F_{ST} . In *Nature Reviews Genetics* (Vol. 10, Issue 9, pp. 639–650). <https://doi.org/10.1038/nrg2611>
- Kania-Gierdziewicz, J., & Pałka, S. (2019). Effect of inbreeding on fertility traits in five dog breeds. *Czech Journal of Animal Science*, 64(3), 118–129. <https://doi.org/10.17221/104/2017-CJAS>
- Kraus, C., Snyder-Mackler, N., & Promislow, D. E. L. (2023). How size and genetic diversity shape lifespan across breeds of purebred dogs. *GeroScience*, 45(2), 627–643. <https://doi.org/10.1007/s11357-022-00653-w>
- Leroy, G., Phocas, F., Hedan, B., Verrier, E., & Rognon, X. (n.d.). *Inbreeding impact on litter size and survival in selected canine breeds 1 2 3*.

- Marelli, S. P., Beccaglia, M., Bagnato, A., & Strillacci, M. G. (2020). Canine fertility: The consequences of selection for special traits. *Reproduction in Domestic Animals*, 55(S2), 4–9. <https://doi.org/10.1111/rda.13586>
- McMillan, K. M., Bielby, J., Williams, C. L., Upjohn, M. M., Casey, R. A., & Christley, R. M. (2024). Longevity of companion dog breeds: those at risk from early death. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-023-50458-w>
- R Core Team. (2024). R: The R Project for Statistical Computing [Computer software]. *R Foundation for Statistical Computing*. <https://www.R-project.org/>
- Strandberg, E., Andersson, L., Emanuelson, U., Bjørnvad, C. R., Ringmark, S., Hedhammar, Å., & Höglund, K. (2023). Heritability and genetic trend of body weight in dogs of different breeds in Sweden. *Journal of Animal Science*, 101. <https://doi.org/10.1093/jas/skad173>
- Yordy, J., Kraus, C., Hayward, J. J., White, M. E., Shannon, L. M., Creevy, K. E., Promislow, D. E. L., & Boyko, A. R. (2020). Body size, inbreeding, and lifespan in domestic dogs. *Conservation Genetics*, 21(1), 137–148. <https://doi.org/10.1007/s10592-019-01240-x>
- Webmatematik. (n.d.). Fordelingsfunktion og frekvensfunktion. Retrieved June 6, 2024, from <https://www.webmatematik.dk/lektioner/matematik-b/statistik/fordelingsfunktion-og-frekvensfunktion>

Appendix

Appendix 1 – Breed information from DKK

A part of the breed information from DKK.

	A	B	C	D	E	F	G	H	I	J	K
1	FCI-gruppe	Race	RAID	Antal hvalpe	Ant.kuld	Ant.basis+	Har Avlsanb.	Ant.Hanner	Ant.Tæver	Antal Importer	I alt
2	FCI-gruppe	X-registreringer	RAID	Antal hvalpe	Ant.kuld	Ant.basis+	Har Avlsanb.	Ant.Hanner	Ant.Tæver	Antal Importer	I alt
3	Total			33374	6557	3410	236	16441	16933	2100	35474
4	'08	Labrador Retriever	'1220	4216	629	398	Ja	2067	2149	105	4321
5	'09	Bichon Havanais	'2500	2331	473	297	Ja	1130	1201	32	2363
6	'01	Schæferhund, stockhåret	'1660	1839	325	135	Ja	903	936	73	1912
7	'08	Golden Retriever	'1110	2240	312	243	Ja	1086	1154	38	2278
8	'09	Coton de Tulear	'2830	853	258	174	Ja	439	414	18	871
9	'09	Cavalier King Charles Spaniel	'1360	973	237	0		469	504	15	988
10	'03	Staffordshire Bull Terrier	'0760	950	195	11	Ja	457	493	34	984
11	'03	Jack Russell Terrier	'3450	572	144	55	Ja	275	297	28	600
12	'08	Engelsk Cocker Spaniel	'0050	577	120	18	Ja	284	293	59	636
13	'02	Berner Sennenhund	'0450	540	118	91	Ja	279	261	26	566
14	'01	Shetland Sheepdog	'0880	399	106	65	Ja	188	211	22	421
15	'02	Dansk/Svensk Gårdhund	'3560	475	104	64	Ja	220	255	10	485
16	'01	Border Collie	'2970	596	101	59	Ja	304	292	54	650
17	'09	Pudel, dværg	'1722	294	95	59	Ja	129	165	17	311
18	'03	West Highland White Terrier	'0850	303	92	45	Ja	139	164	9	312
19	'03	Cairn Terrier	'0040	336	81	55	Ja	167	169	9	345
20	'09	Shih Tzu	'2080	365	80	53	Ja	176	189	24	389
21	'07	Ruhåret Høsehund	'0980	604	75	60	Ja	305	299	20	624
22	'01	Welsh Corgi Cardigan	'0380	459	70	53	Ja	221	238	18	477
23	'09	Tibetansk Spaniel	'2310	246	67	49	Ja	120	126	15	261
24	'09	Chihuahua, langhåret	'2182	161	65	41	Ja	82	79	9	170
25	'01	Australian Shepherd	'3420	459	64	39	Ja	238	221	26	485
26	'07	Kleiner Münsterländer	'1020	446	64	54	Ja	206	240	5	451
27	'05	Finsk Lapphund	'1890	279	60	45	Ja	149	130	15	294
28	'04	Ruhåret Gravhund	'1487	316	60	43	Ja	185	131	15	331
29	'09	Fransk Bulldog	'1010	253	58	0		132	121	24	277
30	'02	Rottweiler	'1470	348	50	2	Ja	184	164	26	374
31	'09	Chihuahua, korthåret	'2181	96	49	31	Ja	46	50	4	100
32	'01	Schæferhund, langstockh. m. underuld	'1661	177	48	24	Ja	85	92	7	184
33	'02	Boxer	'1440	246	47	24	Ja	120	126	23	269
34	'09	Pudel, toy	'1721	113	46	27	Ja	52	61	5	118
35	'03	Border Terrier	'0100	182	45	30	Ja	86	96	9	191
36	'05	Eurasier	'2910	270	45	17	Ja	142	128	6	276
37	'05	Islandsk Fårehund	'2890	219	45	0		109	110	10	229
38	'09	Mops	'2530	161	43	27	Ja	81	80	15	176
39	'08	Flat Coated Retriever	'1210	308	41	21	Ja	142	166	8	316
40	'05	Pomeranian	'0974	93	41	3	Ja	45	48	21	114
41	'08	Engelsk Springer Spaniel	'1250	253	40	22	Ja	146	107	20	273

Appendix 2 – R-coding, linear regression

R-coding for linear regression over average bodyweight and litter size.

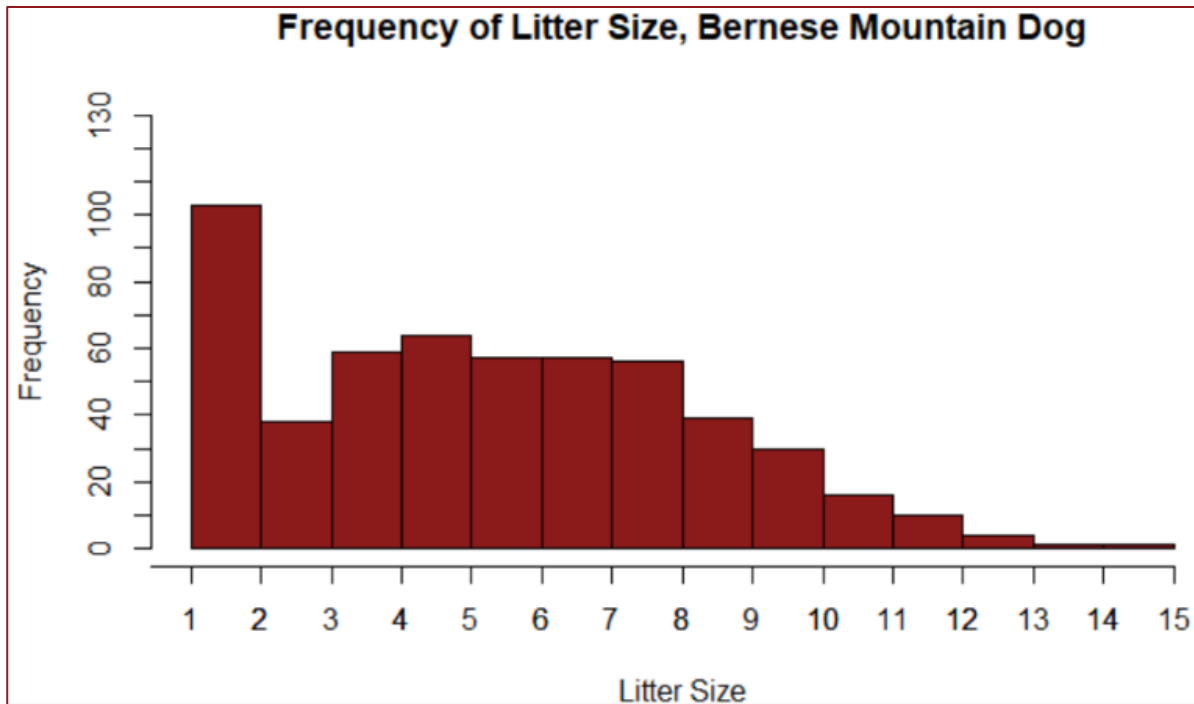
```
```{r}
mod1<-lm(gnm.vægt~kuldstørrelse,data=dkk)
mod2<-lm(log(gnm.vægt)~log(kuldstørrelse),data=dkk)
```

```{r}
plot(dkk$gnm.vægt, dkk$kuldstørrelse, xlab = "gnm.vægt"
, ylab = "kuldstørrelse", main = "Scatterplot", pch = 16)
plot(predict(mod2), rstandard(mod2), xlab = "Prædikterede værdier"
, ylab = "Standardiserede residualer")
abline(h = c(-2, 0, 2), lty = 2)
qqnorm(rstandard(mod2))
abline(0,1)
```

```{r}
plot(dkk$gnm.vægt,dkk$kuldstørrelse)
plot(log(dkk$gnm.vægt),log(dkk$kuldstørrelse))
```

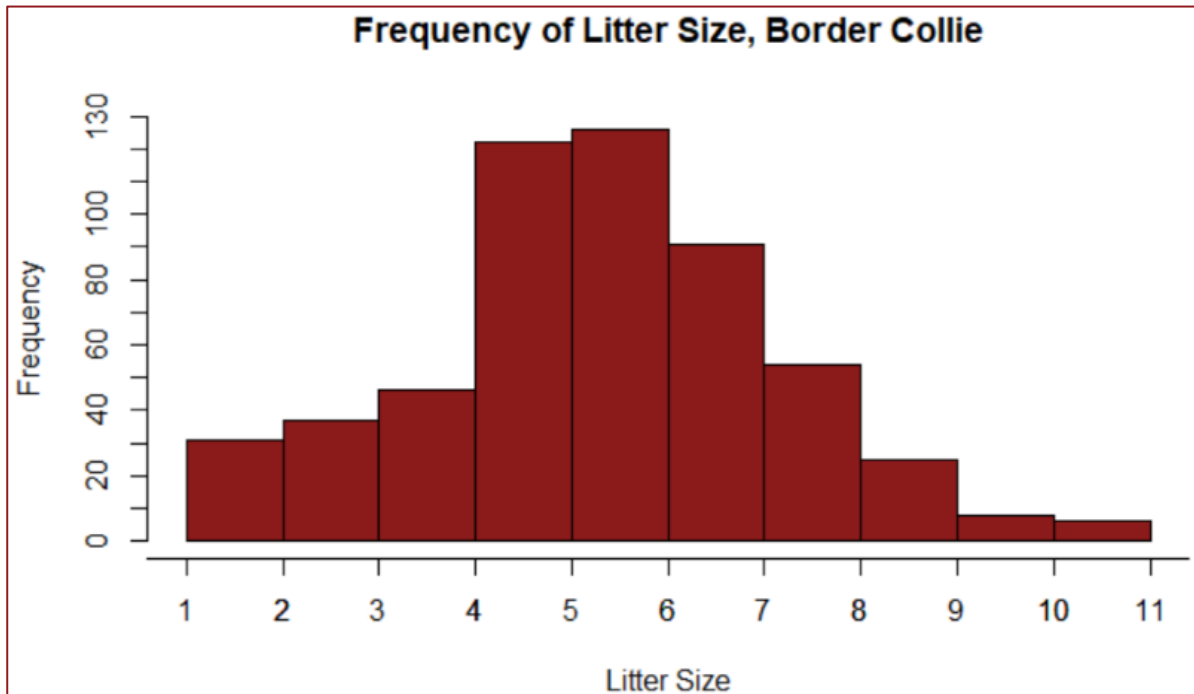
### Appendix 3 – Distribution of litter size for Bernese Mountain Dog

Histogram of the distribution of litter size for the Bernese Mountain Dog with given values: Mean = 5.7, SD = 3.0, SE = 0.1, quartiles (25%, 50%, 75%) as 3, 6, 8, median = 6, Min and Max as 1 and 15.



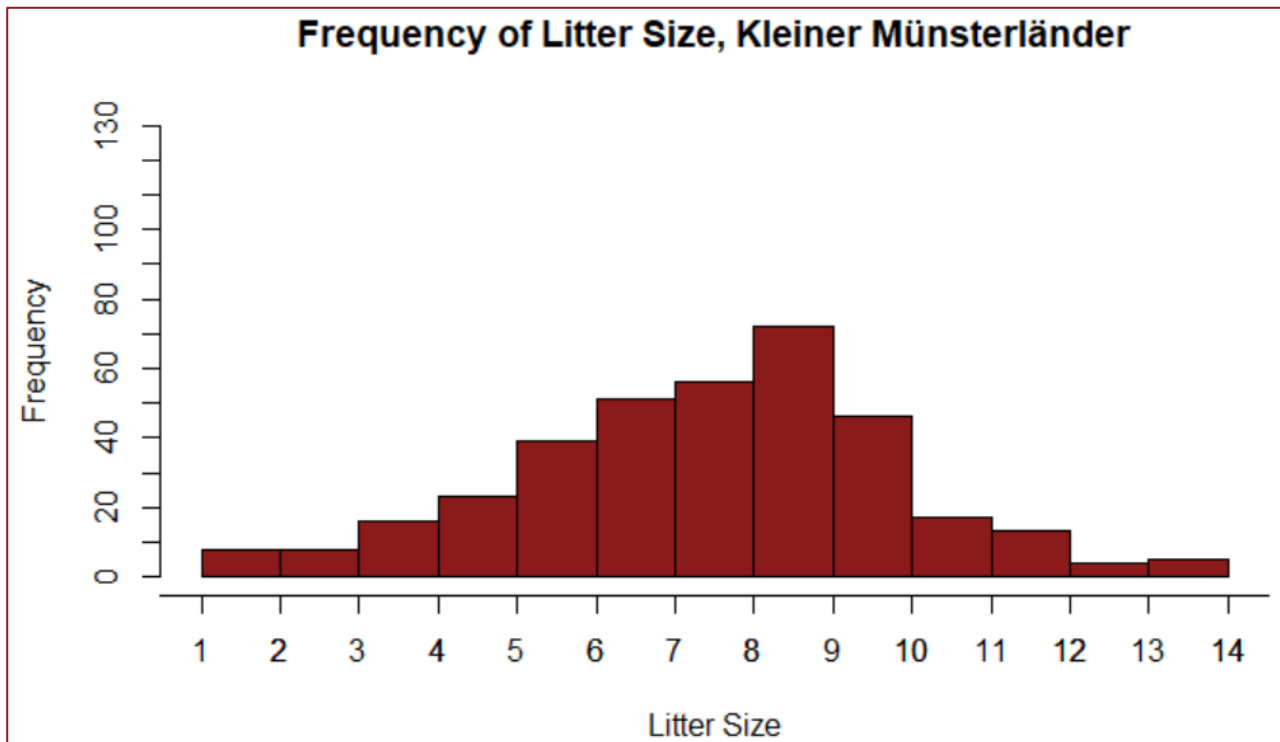
#### Appendix 4 - Distribution of litter size for Border Collie

Histogram of the distribution of litter size for the Border Collie with given values: Mean = 5.8, SD = 1.9, SE = 0.08, quatlilers (25%, 50%, 75%) as 5, 6, 7, median = 6, Min and Max as 1 and 11.



## Appendix 5 - Distribution of litter size for Kleiner Münsterländer

Histogram of the distribution of litter size for the Border Collie with given values: Mean = 7.9, SD = 2.4, SE = 0.1, quantiles (25%, 50%, 75%) as 6, 8, 9, median = 8, Min and Max as 1 and 14.



## Appendix 6 - Distribution of litter size for Rottweiler

Figure 10: Histogram of the distribution of litter size for Border Collies. The given values are as follows: Mean = 7.2, SD = 3.0, SE = 0.1, quartiles (25%, 50%, 75%) = 5, 7, 9, median = 7, minimum = 1, and maximum = 14.

