

Kunskapskontroll 1 i statistik och dataanalys - SELF ASSESSMENT

1. What was the most enjoyable part of the knowledge test?

The most enjoyable part was analyzing the relationships between key variables like price, year, and mileage and visualizing these insights through various plots such as scatter plots, regression lines, and box plots. It was particularly satisfying to observe how luxury brands like BMW and Mercedes consistently stood out in terms of price and variability while identifying the clear negative correlation between mileage and price. Creating these visualizations not only made the data more accessible but also highlighted patterns that would be difficult to identify through raw numbers alone. Furthermore, the process of translating these insights into actionable recommendations was rewarding, as it demonstrated how statistical analysis can provide valuable, real-world insights.

2. What grade do you think you should get and why?

I believe I deserve a grade that reflects the quality and depth of my work, as well as my ability to effectively apply my knowledge in practice through clear and thorough analysis. The analysis is well-structured and covers essential aspects such as price vs. year, price vs. mileage, and brand Influence. Each visualization is paired with clear explanations and meaningful insights that reveal underlying patterns in the data. I applied advanced visualization techniques, such as annotated heatmaps and pair plots, to provide a comprehensive overview of variable relationships. Additionally, I handled data quality challenges effectively and used aggregation methods to simplify complex relationships. My work is thorough, detailed, and demonstrates a strong grasp of both statistical methods and their practical applications. The insights I provided are actionable and relevant, making the analysis valuable for decision-making. The analysis demonstrates a strong commitment to quality, with careful attention to detail and thorough analytical depth.

3. What was the most challenging part of the work and how did you handle it?

The most challenging part was maintaining a balance between detail and clarity. Complex visualizations, such as pair plots and scatter plots, required careful selection of variables to prevent overcrowding and confusion. I addressed this by focusing on the most informative variables, including price, year, mileage, and engine size, while using aggregation techniques, such as grouping by brand, to simplify and enhance clarity. Handling errors, such as key errors in column selection, also posed a challenge. I resolved these by systematically inspecting the dataset using functions like `df.info()` and `df.column()` to identify discrepancies and ensure data accuracy. Moreover, I consistently referred to official documentation and experimented with different visualization parameters to enhance both the accuracy and interpretability of the outputs. This methodical approach allowed me to overcome these challenges while maintaining the quality and coherence of the analysis.