# **IBU 512 Advanced Topics in Database Systems**

Project Proposal

**Topic:** Exploratory Data Analysis on the Billionaire Dataset

**Dataset:** [Billionaires Statistics Dataset (2023)](https://www.kaggle.com/datasets/nelgiriyewithana/billionaires-statistics-dataset/)

**Team Member:** Ajla Korman

**Overview:** This project aims to analyze the dataset from the Billionaire Statistics Dataset. The dataset contains information about the net worth, industries, geographical locations, gender, age, and source of wealth of billionaires worldwide. The first step of this project will involve exploring the dataset using exploratory data analysis (EDA) techniques. The goal is to gain insight into the data and determine which variables are most relevant to the study. Further, the analysis will focus on identifying trends in billionaire wealth distribution, comparing industries with the highest concentration of wealth, and examining geographical patterns to highlight regions or countries with the largest number of billionaires. Additionally, the project will explore demographic variables such as gender and age, as well as analyze how billionaires acquired their wealth (e.g., self-made vs. inherited). The results will be used to draw conclusions about global wealth dynamics and provide valuable insights into the characteristics and distribution of billionaires. For example, the project will highlight which industries dominate billionaire wealth, how wealth varies by region, and whether certain demographic factors, such as age or gender, are associated with higher wealth levels. Overall, this project will offer a comprehensive overview of the billionaire demographic and serve as a foundation for future research.