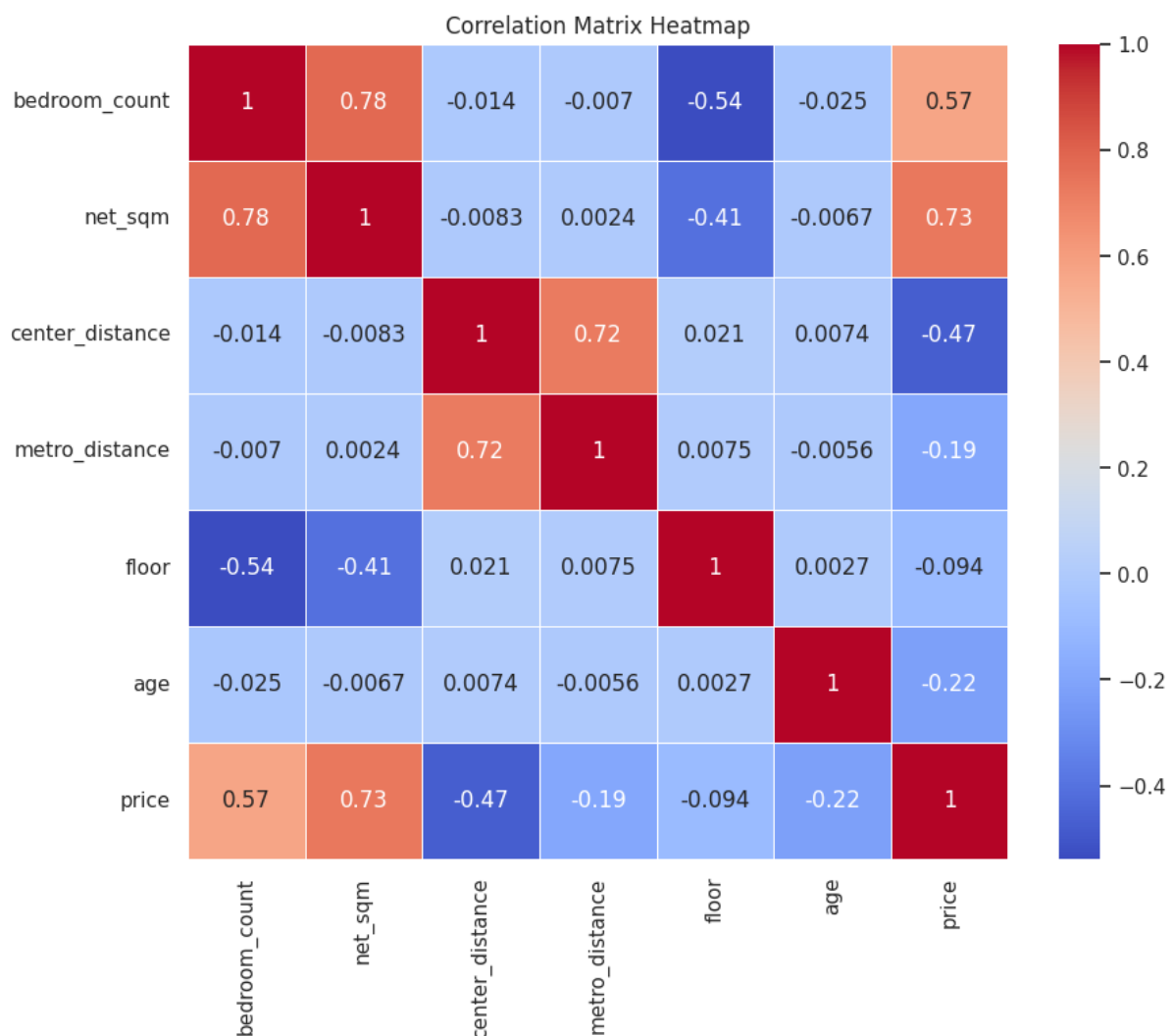# Activity 1 - Data to Insight

Story 1:

Topic: Examination of Housing Price Fluctuations

Data Sources: Housing Market & Prices (Available at: **Kaggle Housing Dataset**)

Description of the story: The attributes in the dataset (like Bedroom Count, Net Square Meters (Net Sqm), Center Distance, Metro Distance, Floor, Age, Price) provide valuable information for potential buyers, renters, and investors, helping them make informed decisions about their housing choices. From the below correlation map, we can see which are important attributes.
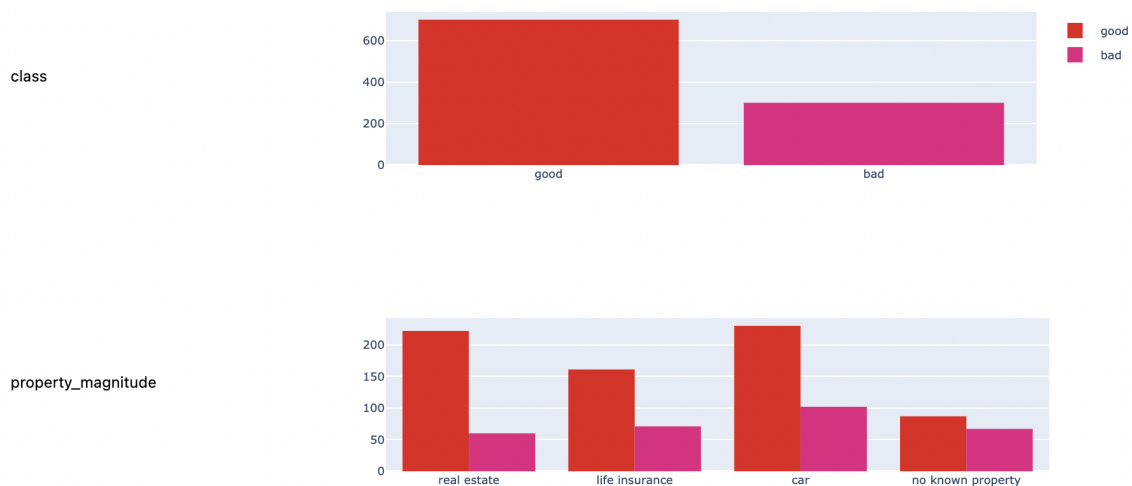

Correlation Matrix Heatmap

Story 2:

Topic: Credit Card Fraud Detection

Data Sources: German Credit dataset(Available at: **Dataset**)

Description of the story: The objective is to leverage this dataset's attributes to develop a robust credit scoring model, enabling the accurate classification of new customers based on their creditworthiness. Such a model holds significant potential for enhancing risk assessment and decision-making processes within the financial industry.
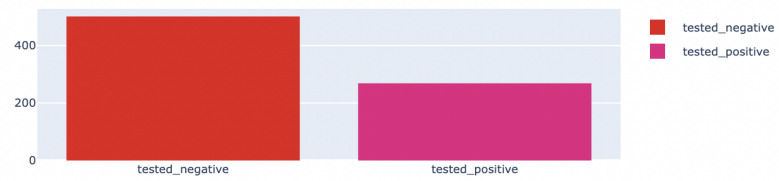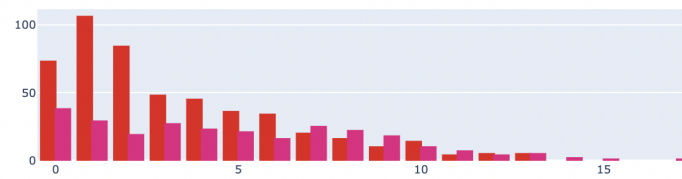

Story 3:

Topic: Diabetes Classification

Data Sources: Diabetes dataset(Available at: **Dataset**)

Description of the story: This dataset presents a valuable opportunity with verity of attributes (like, no. of times pregnant, Diastolic blood pressure, Triceps skin fold thickness, Body mass index, Age, etc.) for a wide range of research studies, including the pivotal task of diabetes classification. The insights and outcomes derived from this study can potentially enhance our understanding of diabetes risk factors and improve diagnostic capabilities in the medical domain.

class

preg

RandomForest Feature Importance