Part B.1

Link to colab

https://colab.research.google.com/drive/14GP2EMjTBCMyF_TyEabGcNNOgEWDV2dA#scrollTo=CmsJKTELLKga

GitHub Repository:

https://github.com/BinxuanWu/CSI4142-Project

Part B.2

Model	Accuracy	Precision	Recall	Time
Decision Tree	0.846402	0.846577	0.846402	6.702s
Gradient Boosting	0.850605	0.849783	0.850605	454.697s
Random Forest	0.879274	0.878462	0.879274	68.428s

By comparing the performance of the three different models, we can identify the strengths and weaknesses of each. In our case, the algorithm of random forest has a higher accuracy and precision. Also, it takes a decent time to construct. It can help us to understand which features are the most important(a higher weight), for predicting housing prices in Beijing.

By comparing the predicted price range and the actual price, we can identify outliers and investigate potential reasons for the discrepancies. This can help us to understand which factors are causing the biggest discrepancies and how we can improve the accuracy of our models.

By analyzing housing prices across different regions of Beijing, we can gain insights into the spatial patterns of housing prices and the factors that drive these patterns. This can help us to understand how location affects housing prices, in our case district and cid. This can help us to make predictions about future housing price trends and adjust our models accordingly.

By examining the coefficients of each feature in the different models, we can gain insights into which features are most important for predicting housing prices. This can help us to understand what factors drive housing prices in Beijing and how they interact with each other.

The model has helped us to understand the impact of property features on housing prices, variable such as location, elevator and subway. By analyzing the results, we can identify the property features that have the greatest impact on housing prices.