Holmusk Assignment Report Draft 1

# Introduction

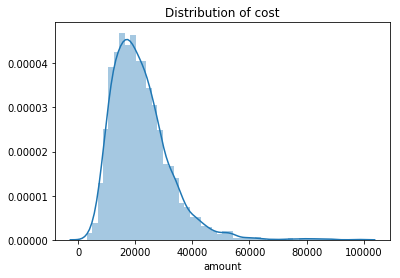
# Data –

## Preprocessing

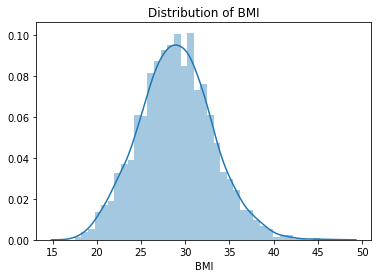
Repeated Patient IDs – Handling strategy? 🡪 Remove all?

## Exploratory Data Analysis (Univariate & Multivariate)

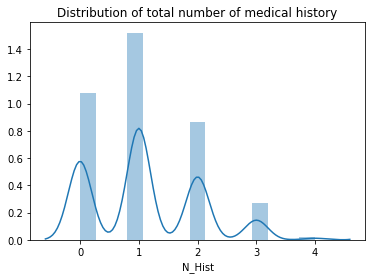
1. Distribution of cost

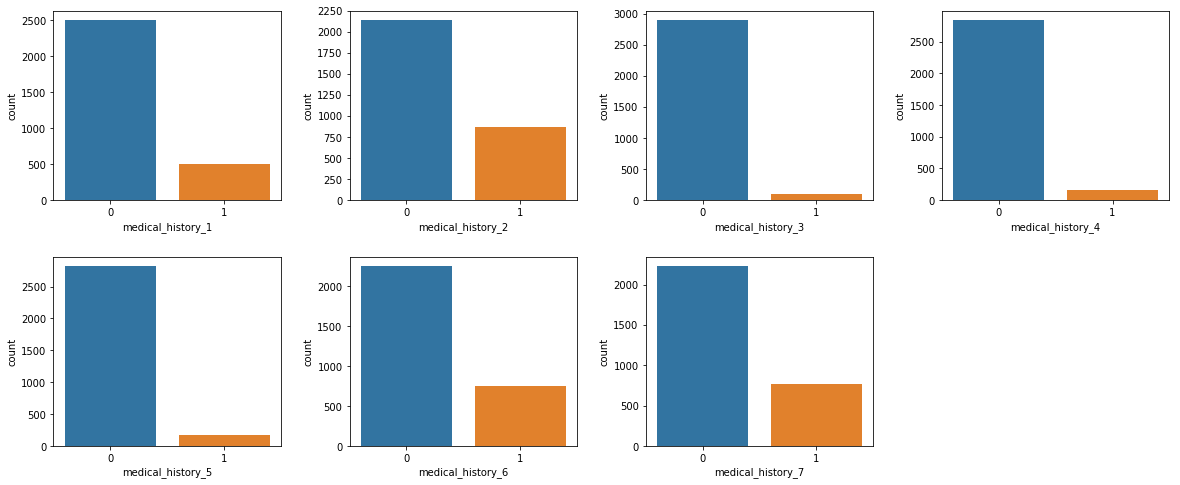


1. Distribution between gender
2. Distribution between races
3. Distribution between age
4. Distribution between BMI

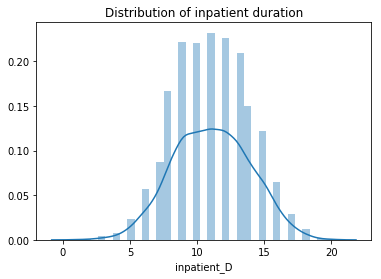


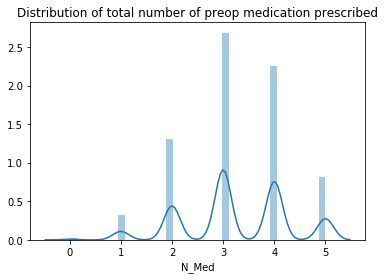
1. Distribution of lab tests
2. Distribution of medical history availability

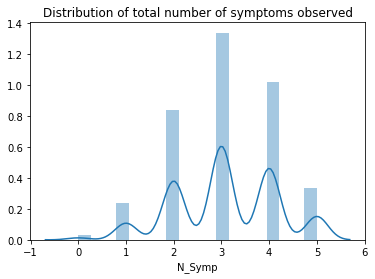




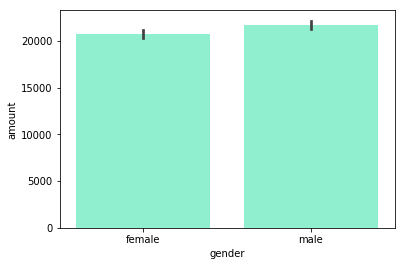
1. Distribution of inpatient duration

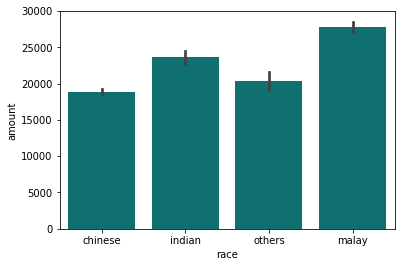


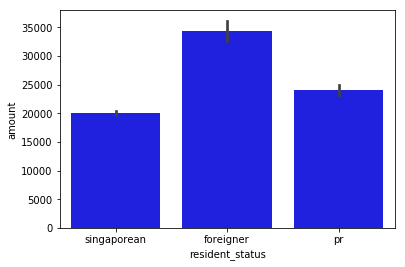
1. Distribution of pre-operation medication
2. Distribution of total number of specific symptoms exhibited



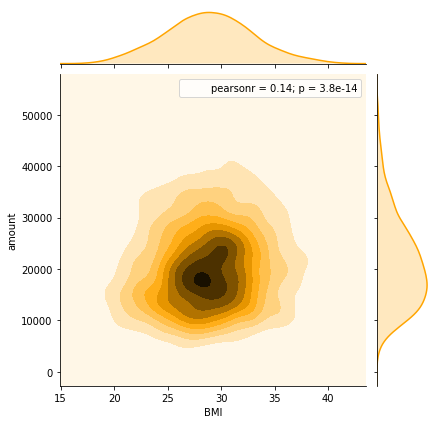
1. 

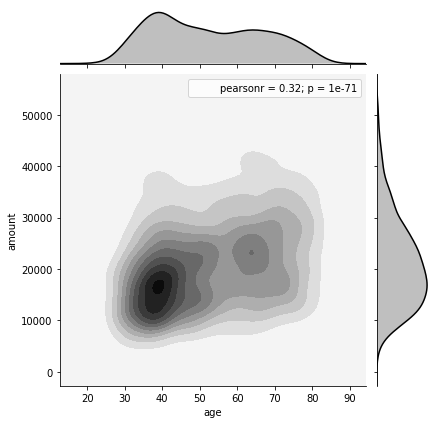


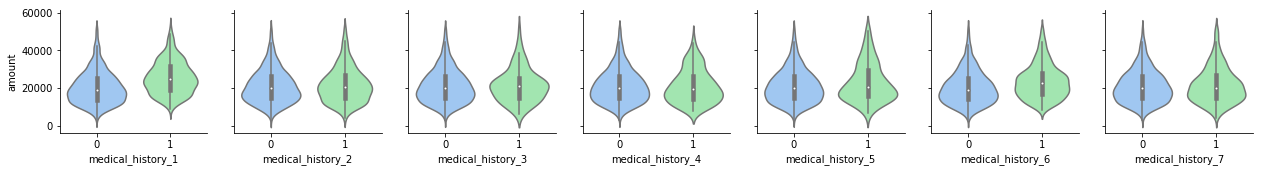


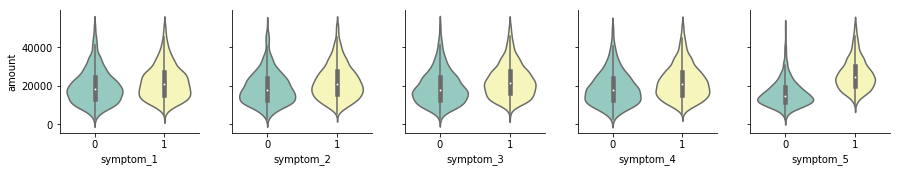


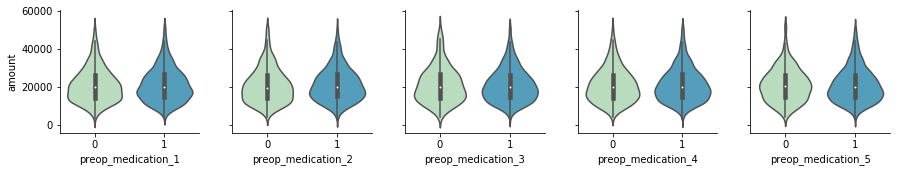


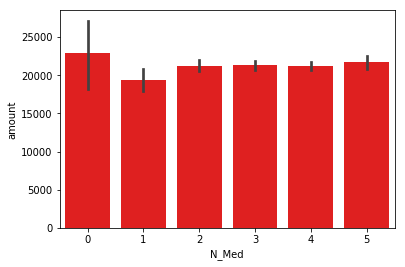


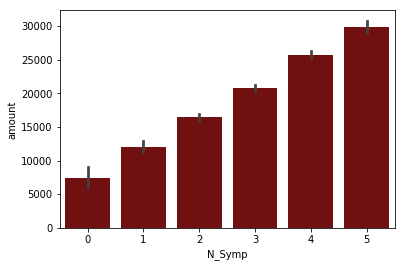


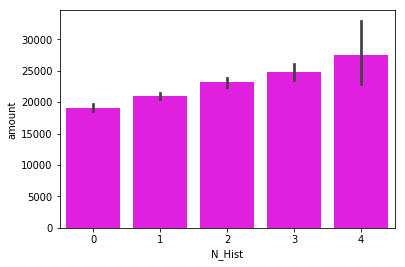


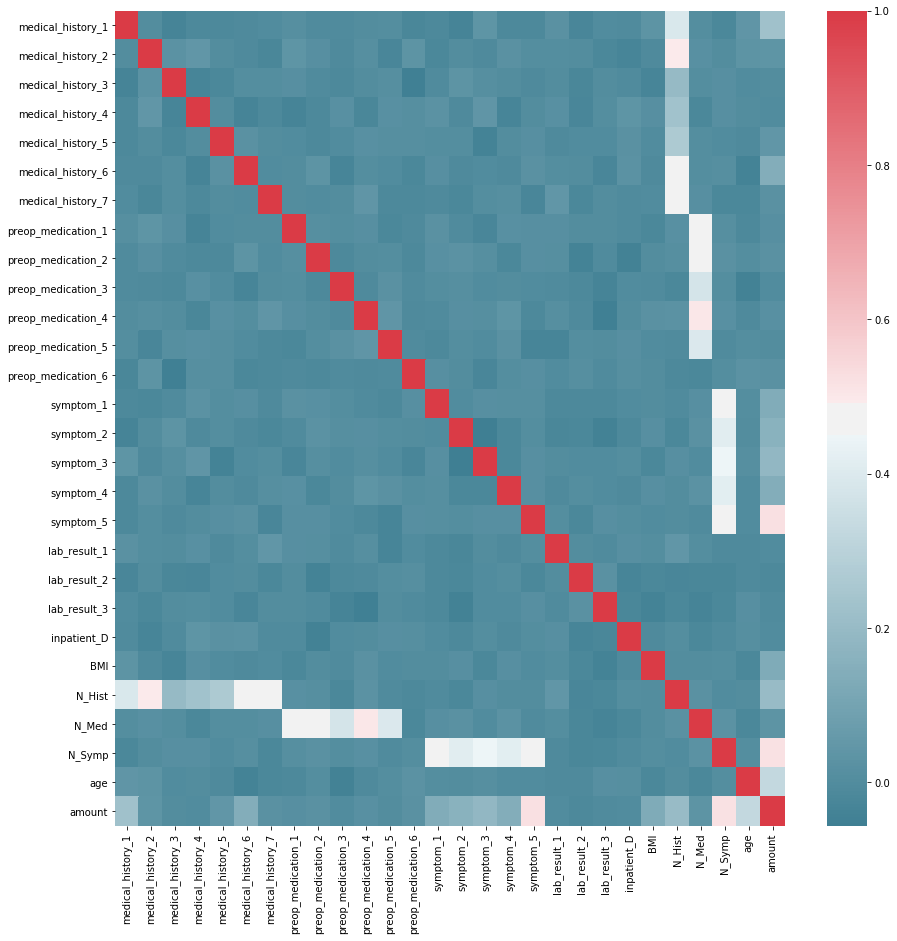


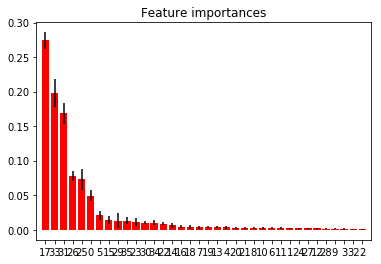












# Method

Random Tr

Cross-validation strategy

Variable Importance

# Results

# Discussion

Statistical Significance vs Practical Significance

Existing Literature

# Conclusion

# Appendix