

Stat149: Final Project Report

Karina Huang, Nicholas Stern, Phoebe Wong

4/18/2019

1 Background

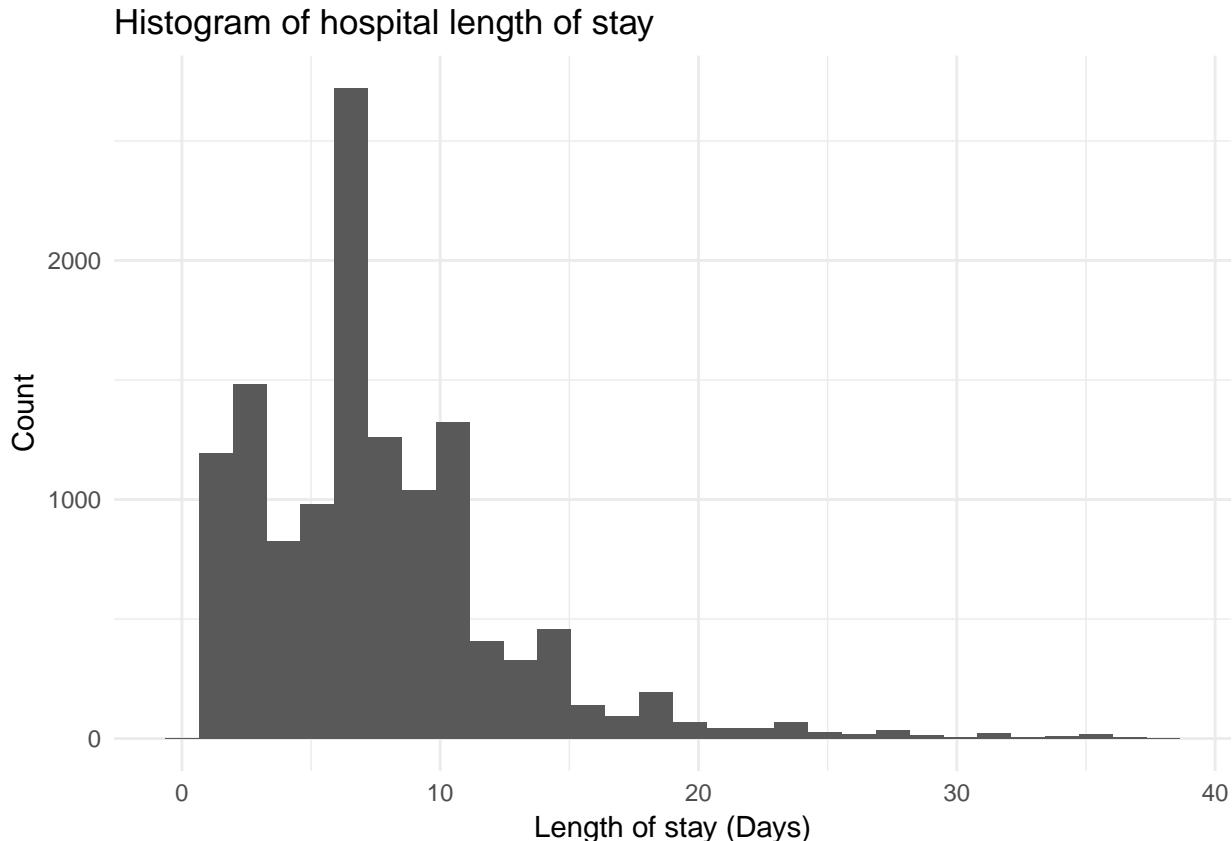
- Description of the problem and its importance
- Description of data
- Goals of project

2 Exploratory Analyses

- Key features of the data that will be relevant for modeling
- Any data preprocessing?

2.1 EDA

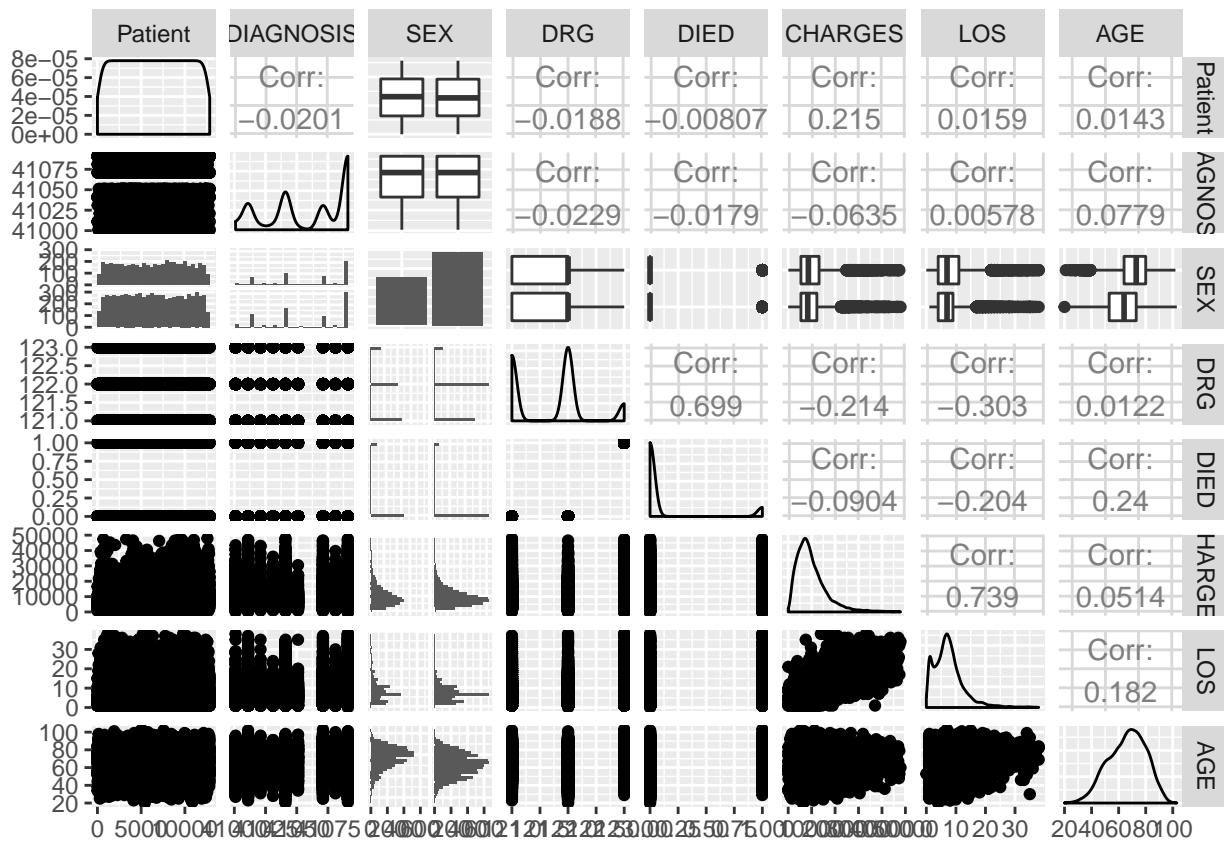
2.1.1 Response variable - Hospital length of stay



We can see that length of stay (LOS) spans from 0 to 38 days, with a median of 7 days.

```
summary(data$LOS)
```

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
##	0.000	4.000	7.000	7.569	10.000	38.000



3 Model

- Main models considered — Gamma (right-skewed), inverse-Gaussian (right-skewed), negative binomial (# of trials), Poisson (discrete count)
- Evaluate model based on Log-likelihood between types of model
- Decisions to consider alternative models

4 Conclusion

- Substantive conclusions that would be important for a non-statistician
- Report any pragmatic advice/conclusions result from the analysis

5 Discussion

- Critical evaluation of your overall approach
- What aspects of the modeling attempts did you expect would substantially improve inference but did not achieve the desired outcome?

- What limitations can you offer about your results or the process that led to your results?