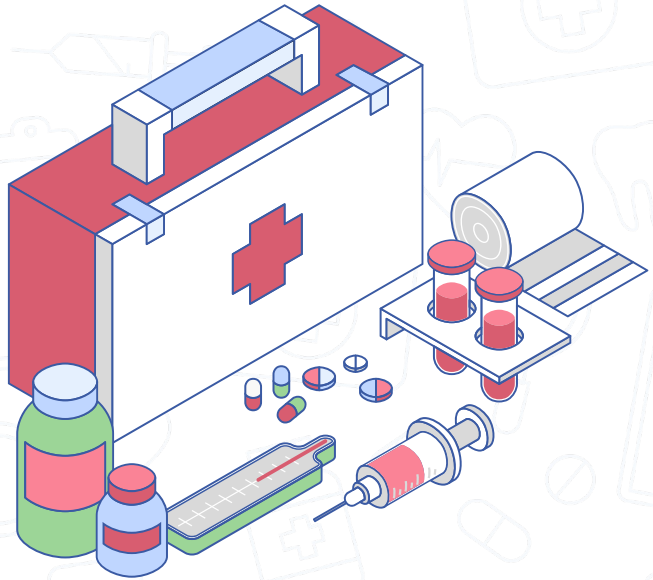
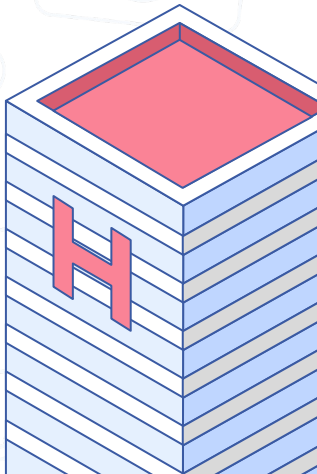


HARMONY HEALTHCARE

FEATURE SELECTION

Hibah Arshad
Samantha Nadler
Thomas Walsh



ACKNOWLEDGEMENTS

DR. ANGEL PINEDA

for helping manage our classwide final project.

DR. LAURIE WARD

for working so hard with us and collaborating with Harmony Healthcare.

BART, DANIEL, EDDIE, JAMARI, AND JUSTINE

for providing feedback on our work thus far and helping on the other stages of the project.

UNITED STATES EMERGENCY ROOM TRENDS

- 16.3% of emergency room patients were resulted in admission
- In 2018, the average rate of alcohol related ER admissions was 163.4 visits per 10,000 visits (1.63%)
- In 2020, that number increased to 184.8 (1.85%)

1) Pines, Jesse M., et al. “Variation in Emergency Department Admission Rates across the United States.” Medical Care Research and Review : MCRR, vol. 70, no. 2, 2013, pp. 218–31, https://doi.org/10.1177/1077558712470565.

2) Marissa B. Esser, Nimi Idaikkadar, Aaron Kite-Powell, Craig Thomas, Kurt J. Greenlund, Trends in emergency department visits related to acute alcohol consumption before and during the COVID-19 pandemic in the United States, 2018–2020, Drug and Alcohol Dependence Reports, Volume 3, 2022, 100049, ISSN 2772-7246, https://doi.org/10.1016/j.dadr.2022.100049.

EMERGENCY ROOM READMISSION TRENDS

10.5%

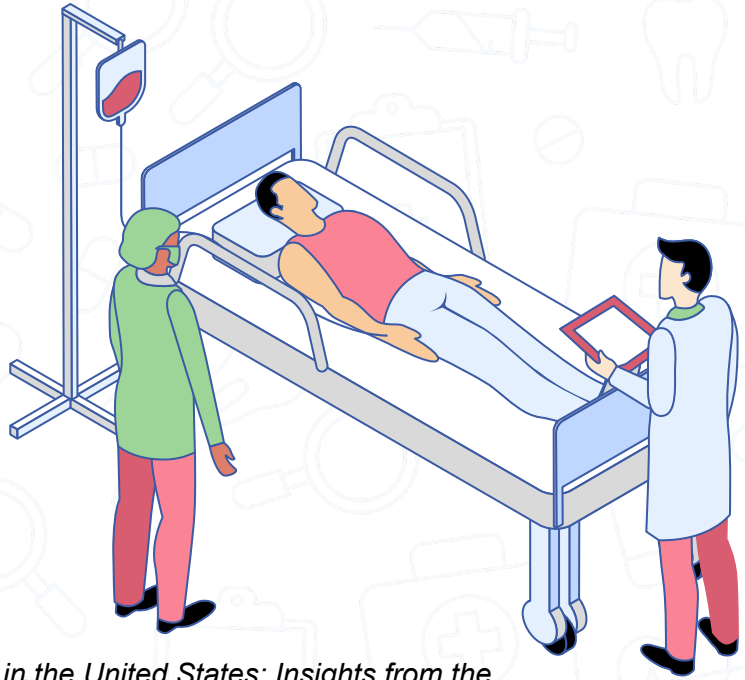
percent of patients readmit themselves to an emergency room within 30 days.

45+

Readmission within 30 days is more common in patients ages 45 and older.

MOST COMMON FACTORS

Heart failure and septicaemia were the most frequent factors causing readmission.

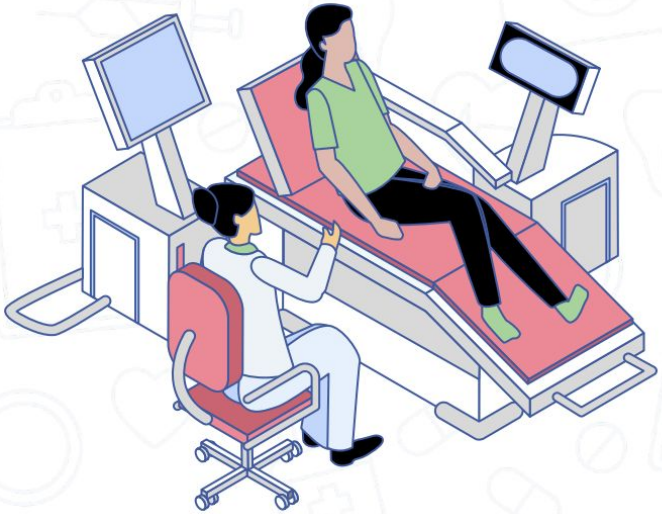


[1] Fonarow, Gregg C, et al. *All-Cause Unplanned Readmissions in the United States: Insights from the Nationwide Readmission Database*, 11 Oct. 2021, doi.org/10.1111/imj.15581.

MOST COMMON READMISSION CAUSES[1]

- Respiratory-related (22%)
- Nervous-related (17.3%)
- Cardiovascular-related (16.7%)
- Digestive-related (11.5%)
- Covid-19 (9.3%)

[1] Glynn, Nigel et al. “Emergency medical readmission: long-term trends and impact on mortality.” *Clinical medicine (London, England)* vol. 11,2 (2011): 114-8. doi:10.7861/clinmedicine.11-2-114
2) Abolfazl Akbari, Amirhossein Fathabadi, Mahya Razmi, Ahmadreza Zarifian, Mahdi Amiri, Alireza Ghodsi, Elnaz Vafadar Moradi, Characteristics, risk factors, and outcomes associated with readmission in COVID-19 patients: A systematic review and meta-analysis, The American Journal of Emergency Medicine, Volume 52, 2022, Pages 166-173, ISSN 0735-6757, https://doi.org/10.1016/j.ajem.2021.12.012.



RESEARCH QUESTION

What factors contribute the greatest to a patient getting readmitted to the emergency room?

- Feature engineering
- Build models to test variable correlation

MODELS WE CONSIDERED USING

LASSO METHOD

- Feature selection and regularization performed simultaneously

LOGISTIC REGRESSION

- Models log-odds of events as linear combinations of variables

XGBOOST

- Iteratively builds decision trees to determine an average prediction

MODEL WE DID USE

LASSO METHOD

- Feature selection and regularization performed simultaneously

LOGISTIC REGRESSION

- Models log-odds of events as linear combinations of variables

XGBOOST

- Iteratively builds decision trees to determine an average prediction

TEAM I'S DATA

Dataset cleanedDatasetV2.csv has no missing values:

- Categorical data all is lowercase to standardize text, replaced missing values with mode
- Numerical to binary to make predictions, replaces missing with mean
- Labels were converted to “Yes” or “No” to address imbalanced dataset

LASSO REGRESSION

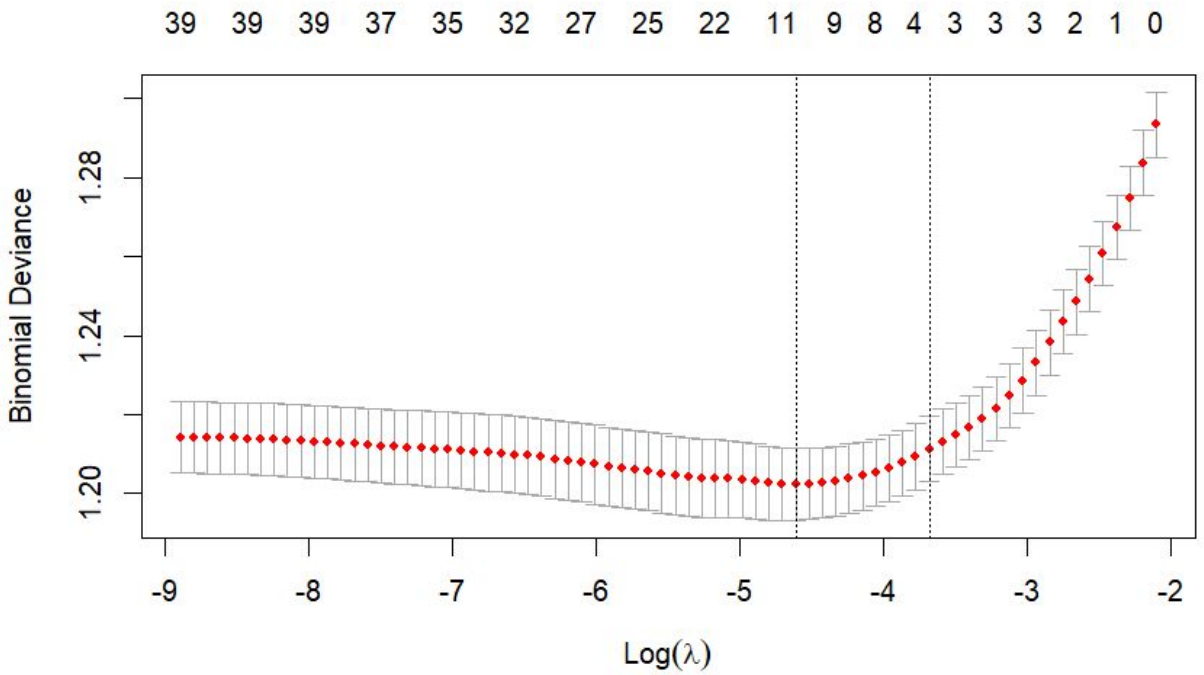
(Least Absolute Shrinkage and Selection Operator)

```
y ← HH_Data$Admission
x ← HH_Data %>%
  select(where(is.numeric), -Admission) %>%
  as.matrix()

library(glmnet)
set.seed(42)

lasso_model ← cv.glmnet(x, y, alpha = 1, family = "binomial")
best_lambda ← lasso_model$lambda.min
cat("Best lambda (from cross-validation): ", best_lambda, "\n")
plot(lasso_model)
```

LASSO REGRESSION



LASSO REGRESSION

- Patient.HCC.Risk.Total.Risk: -0.0446898215
- Active.Medications: 0.0443919593
- Primary.Care.Encounter.Count: 0.0331995417
- SDOH.Assessment.Count: 0.0146984831
- Patient.Appointment.No.Show.Rate..: 0.0146875578
- Depression.Screening.Count.Past.Yr: 0.0086325980
- eGFR.Result: 0.0067602679
- Most.Recent.BMI.Value: 0.0037749693
- UDS.Qualifying.Encounter.Count: 0.0032103970
- COVID.19.Immunization.Code: -0.0006886961
- Fasting.Glucose.Test.Result: -0.0006487510

FUTURE STEPS / LIMITATIONS

USE OTHER PREDICTIVE MODELS

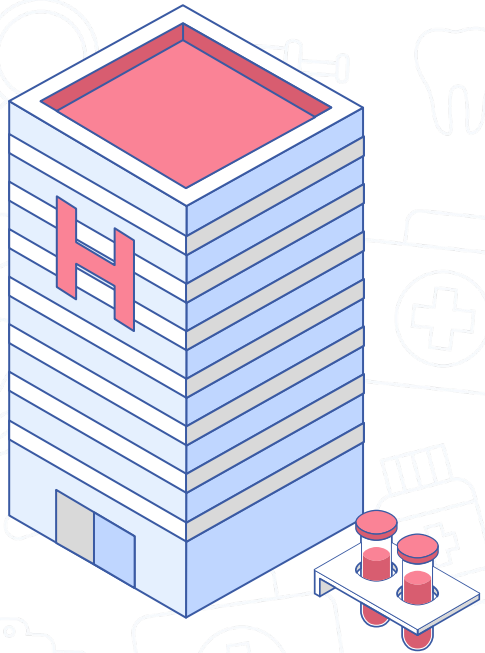
- Consider using logistic regression and/or XGBoost

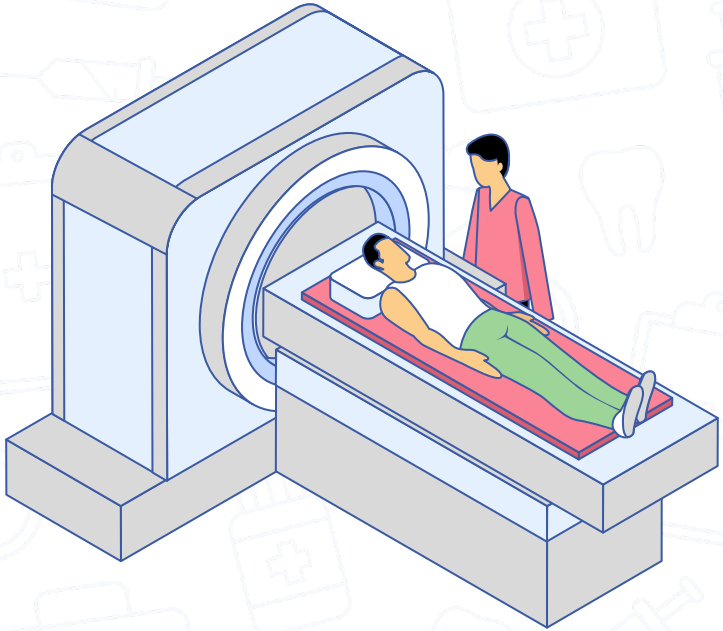
CONSIDER MORE RECENT DATA

- Dataset only contained one week's worth of data

GET FIXED AGES

- Not on us, but age column got corrupted and produces ages all ending in "1900"





THANK YOU!

DOES ANYONE HAVE ANY QUESTIONS?