Data Science Solution in Improving the Occupancy of the Nursing Facilities

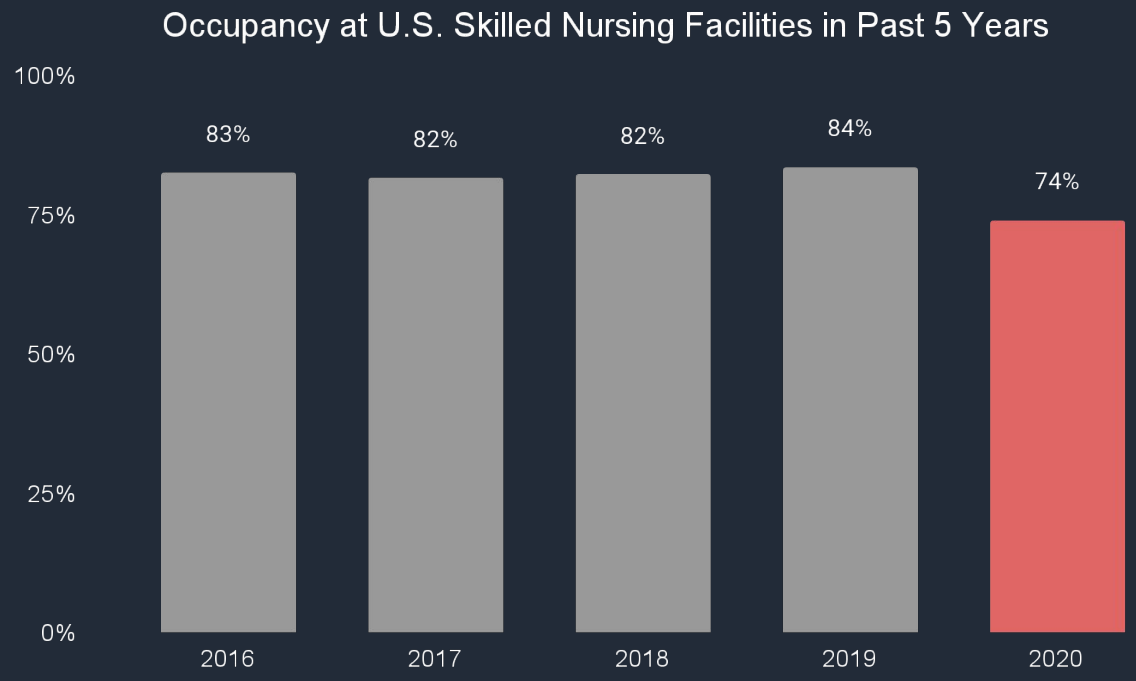
Chien Yuan Chang
10/15/2021

OUR TEAM

We provide data science, machine learning and AI solutions for our clients to solve their problems

What Happened

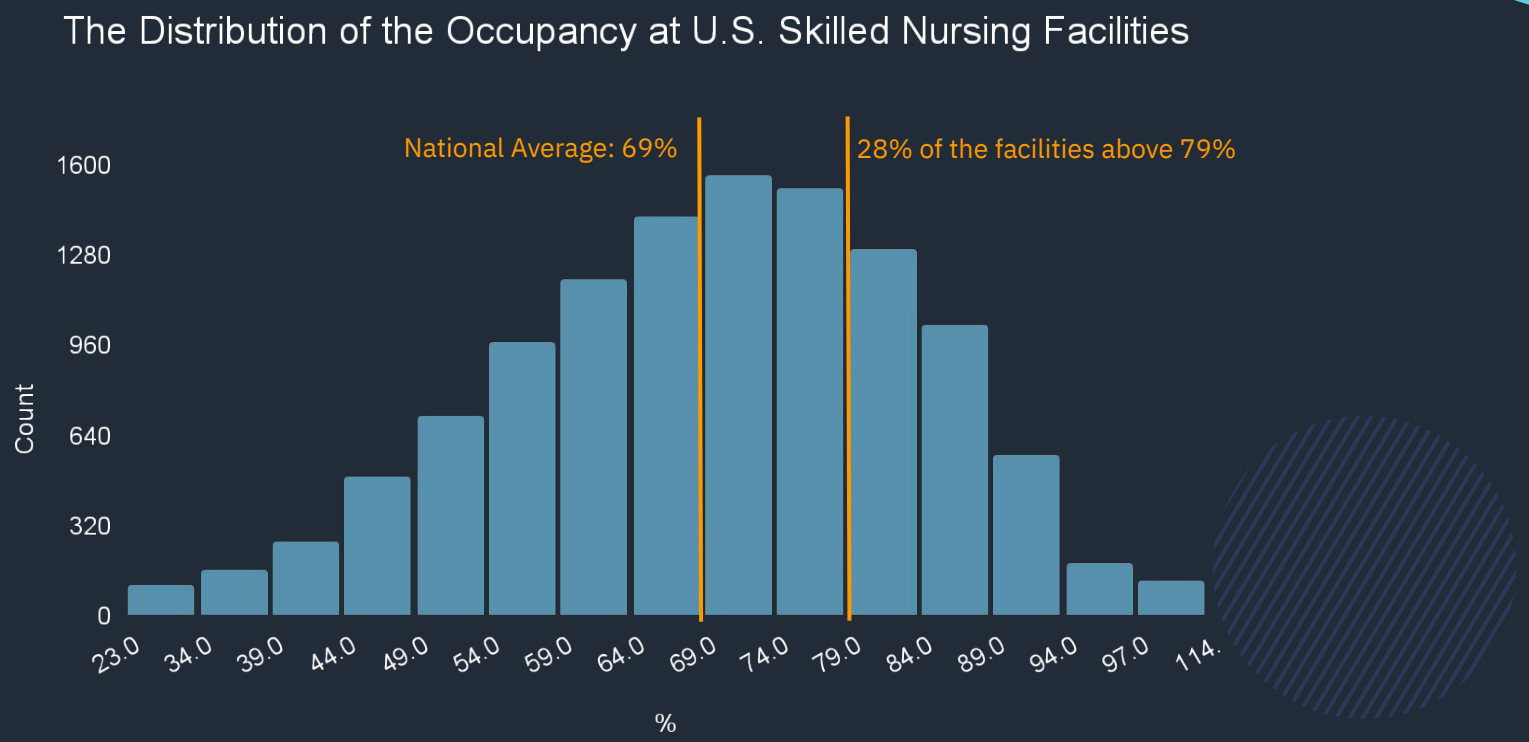
Huge Drop in 2020 Due to the Pandemic



*Data source: National Investment Center

The Problem

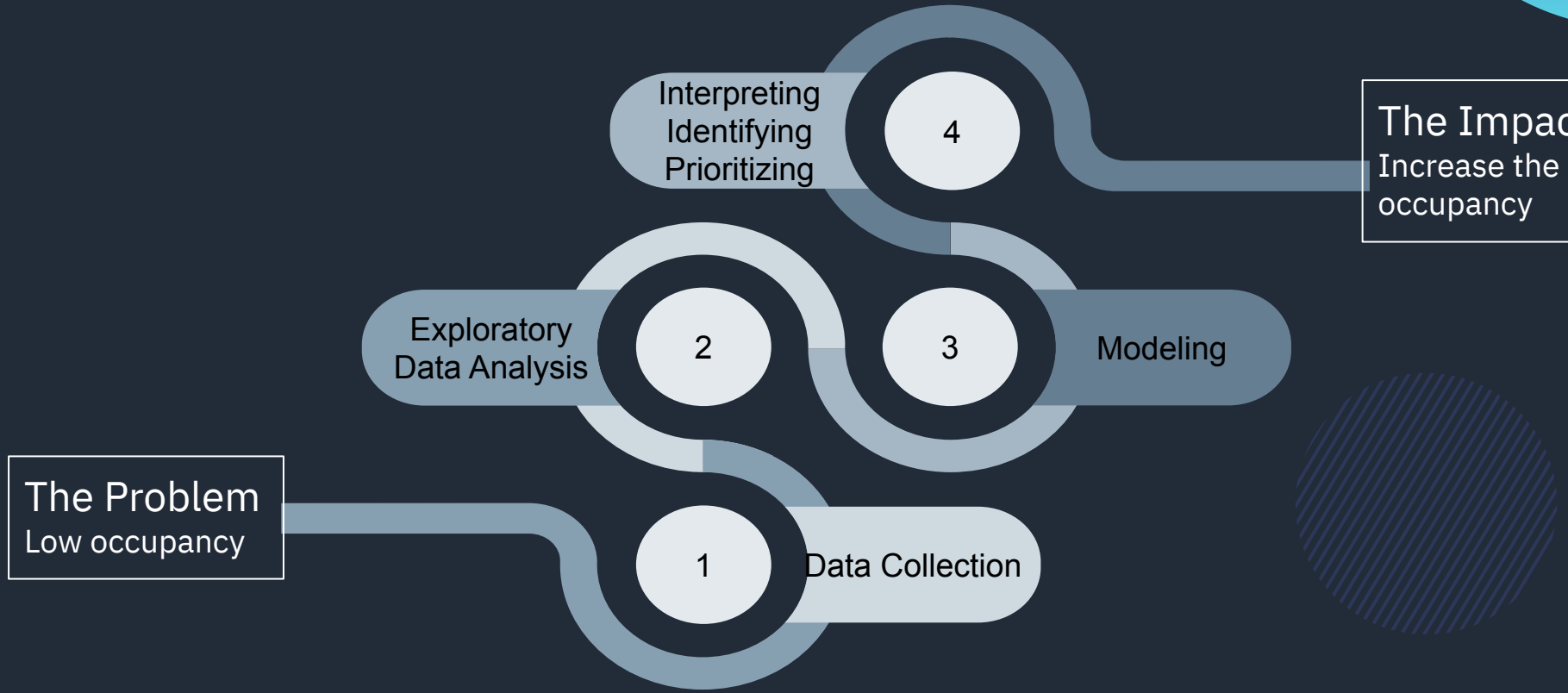
Nursing facilities with low occupancy



*Data source: Centers for Medicare & Medicaid Services

Our Data Science Solutions

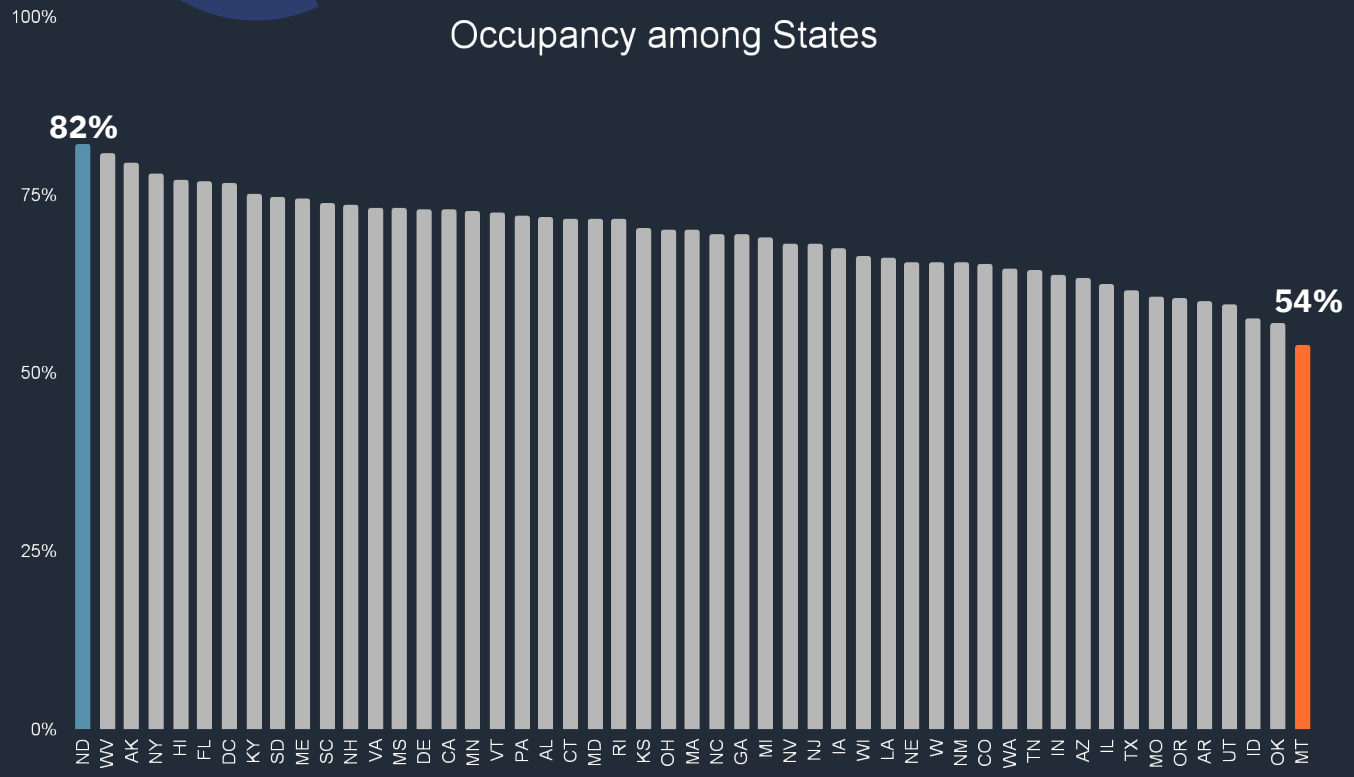
A Regression Model to interpret the factors of the occupancy



Preliminary Analysis - Dataset and Tools

- Data
  - “Provider Information” dataset
    - 15,265 nursing facilities with 88 columns
    - From CMS.gov (the Centers for Medicare & Medicaid Services)
      - 21 datasets of nursing homes including rehab services
- Tools
  - Excel
    - Data cleaning and analysis
  - Tableau
    - Map visualization
  - Google Slides
    - Visualizations and presentation
  - Python
    - Baseline regression model

Preliminary Analysis - Area

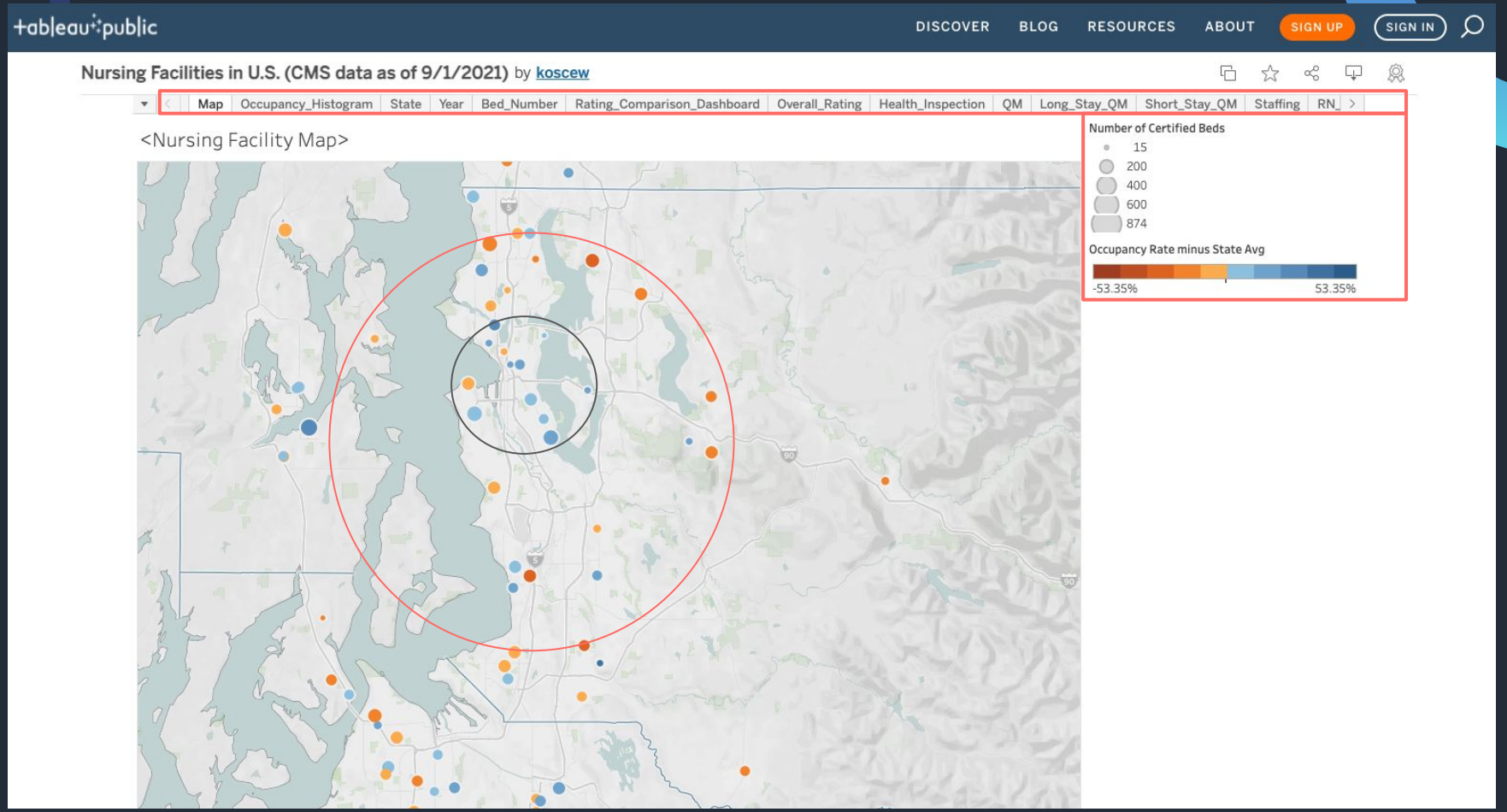


*Data source: Centers for Medicare & Medicaid Services

Control regional effect:

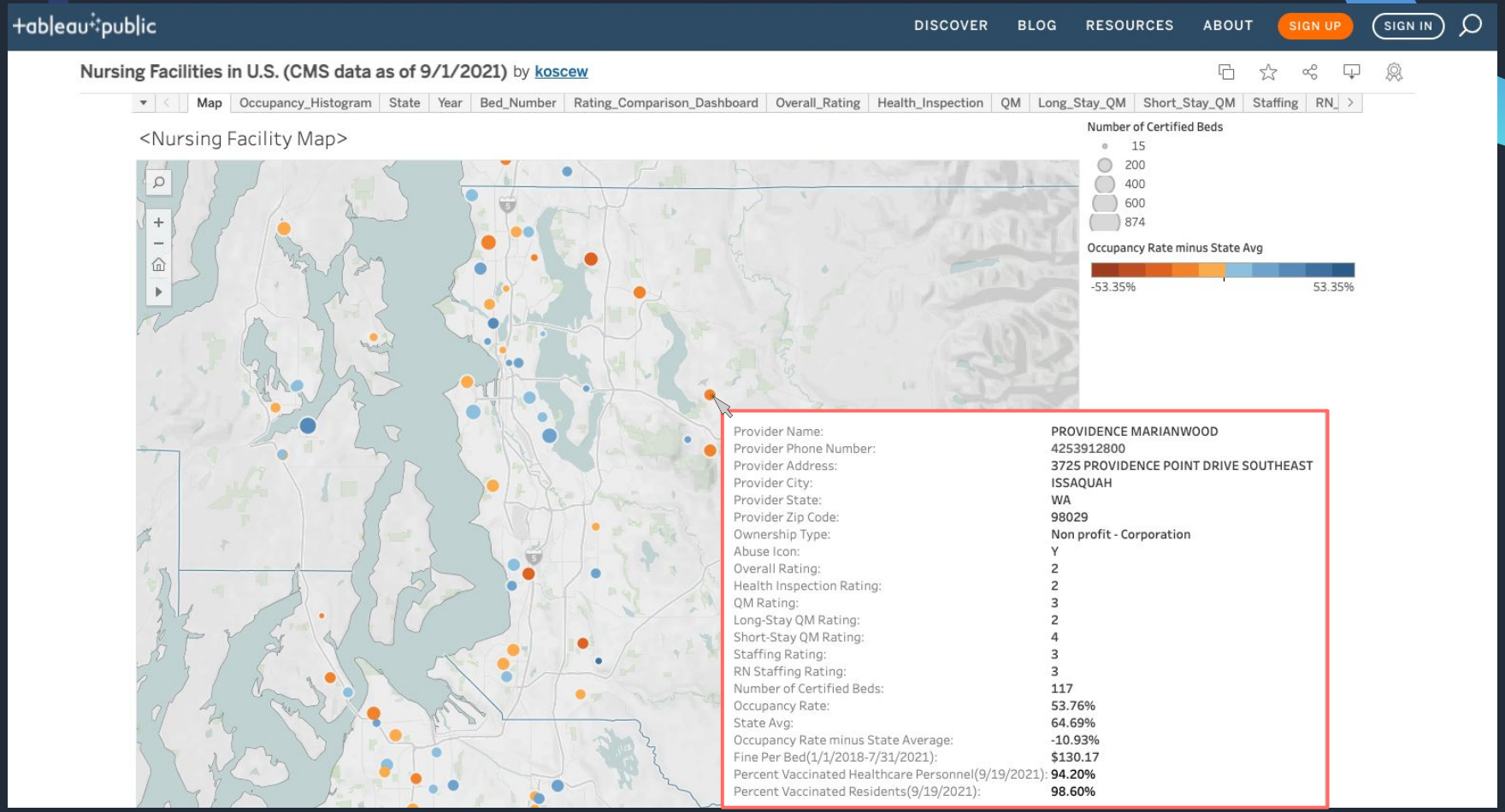
Use the difference between each nursing home and its state average occupancy as the target for EDA

Preliminary Analysis - Area



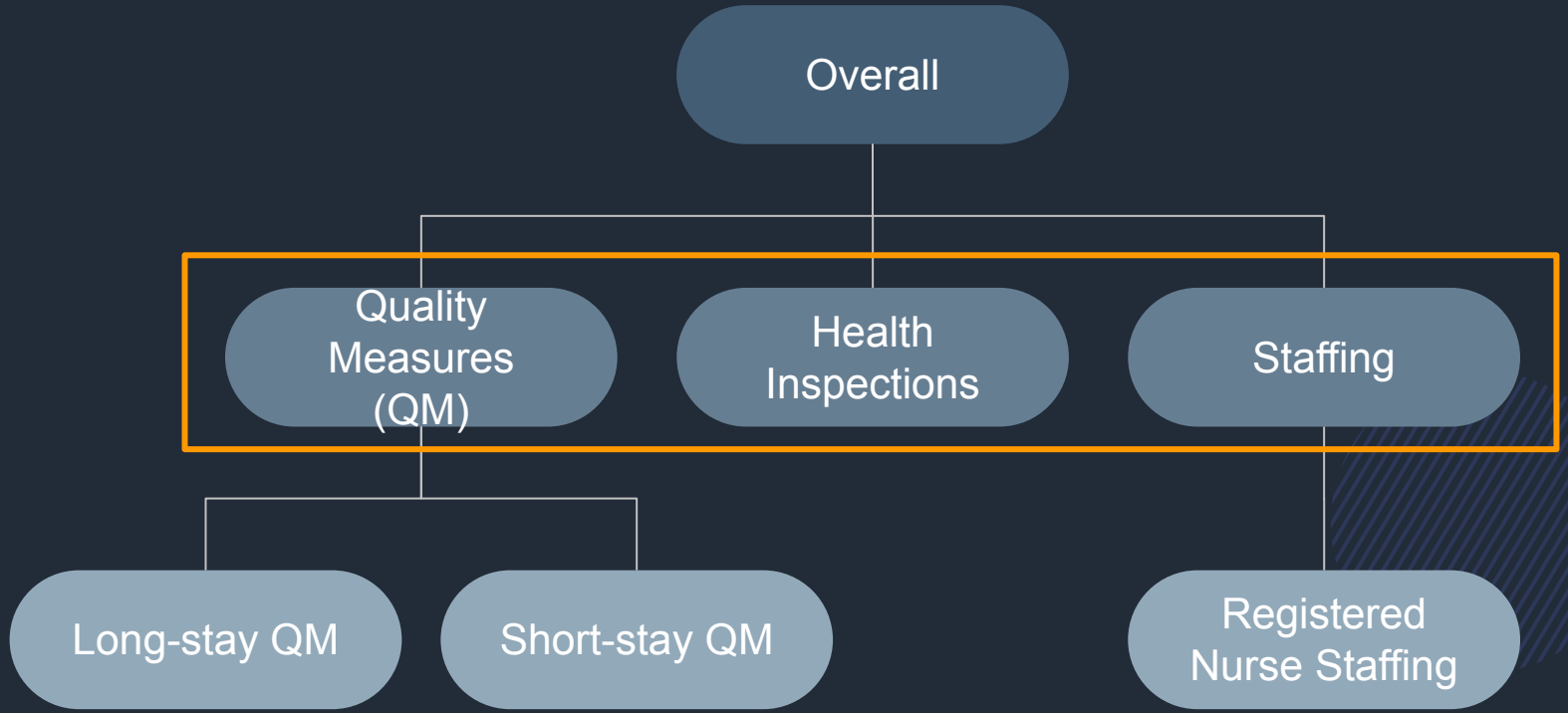
*Data source: Centers for Medicare & Medicaid Services

Preliminary Analysis - Area

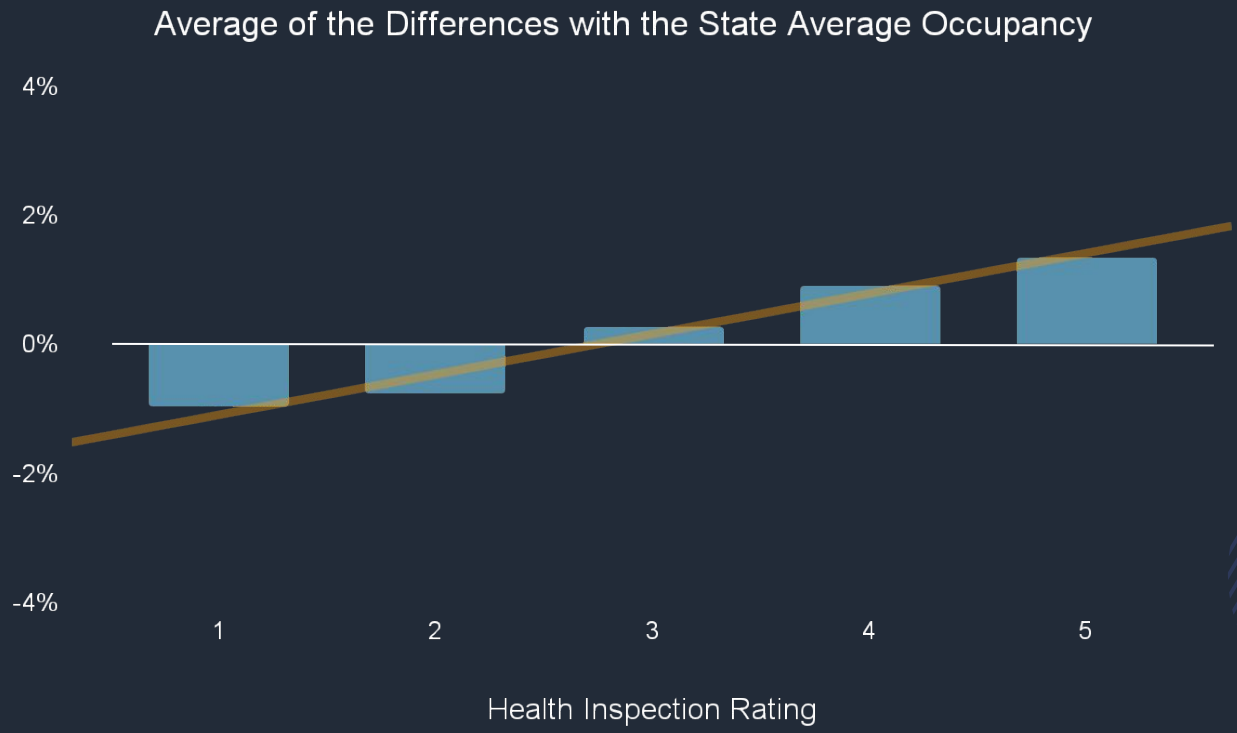


*Data source: Centers for Medicare & Medicaid Services

Preliminary Analysis - Ratings

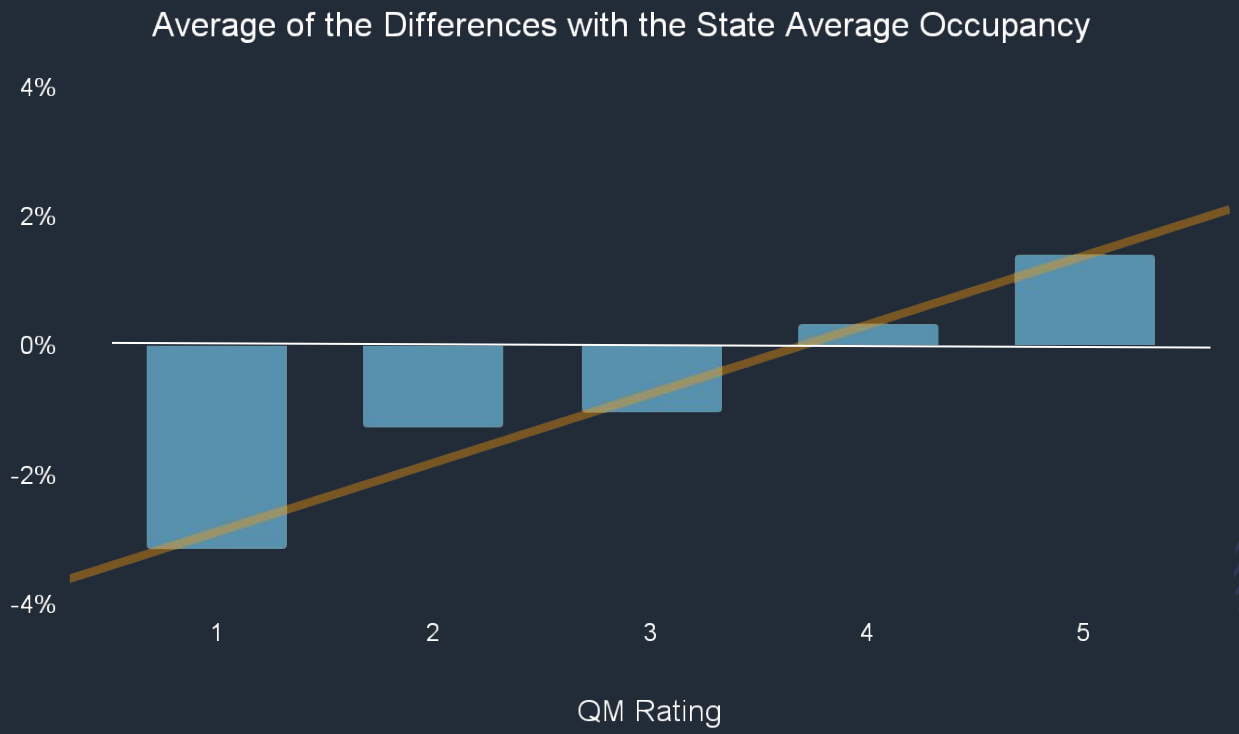


Preliminary Analysis - Health Inspection Rating



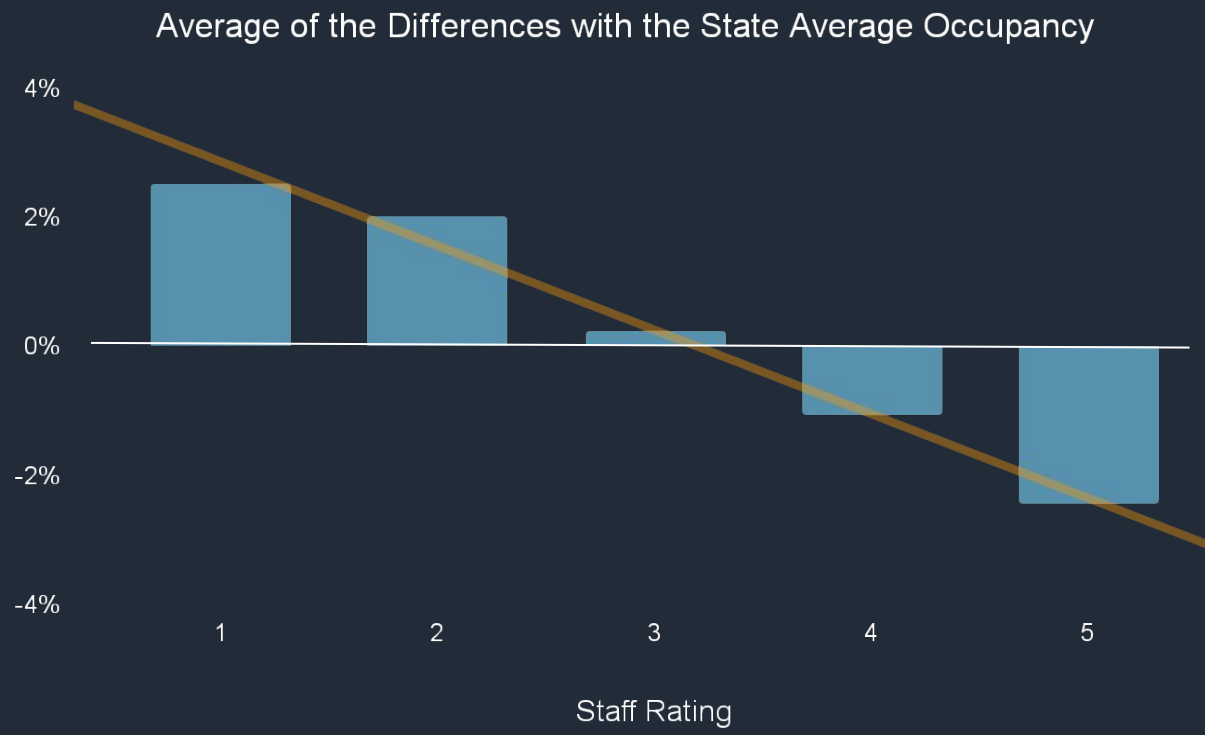
*Data source: Centers for Medicare & Medicaid Services

Preliminary Analysis - QM Rating



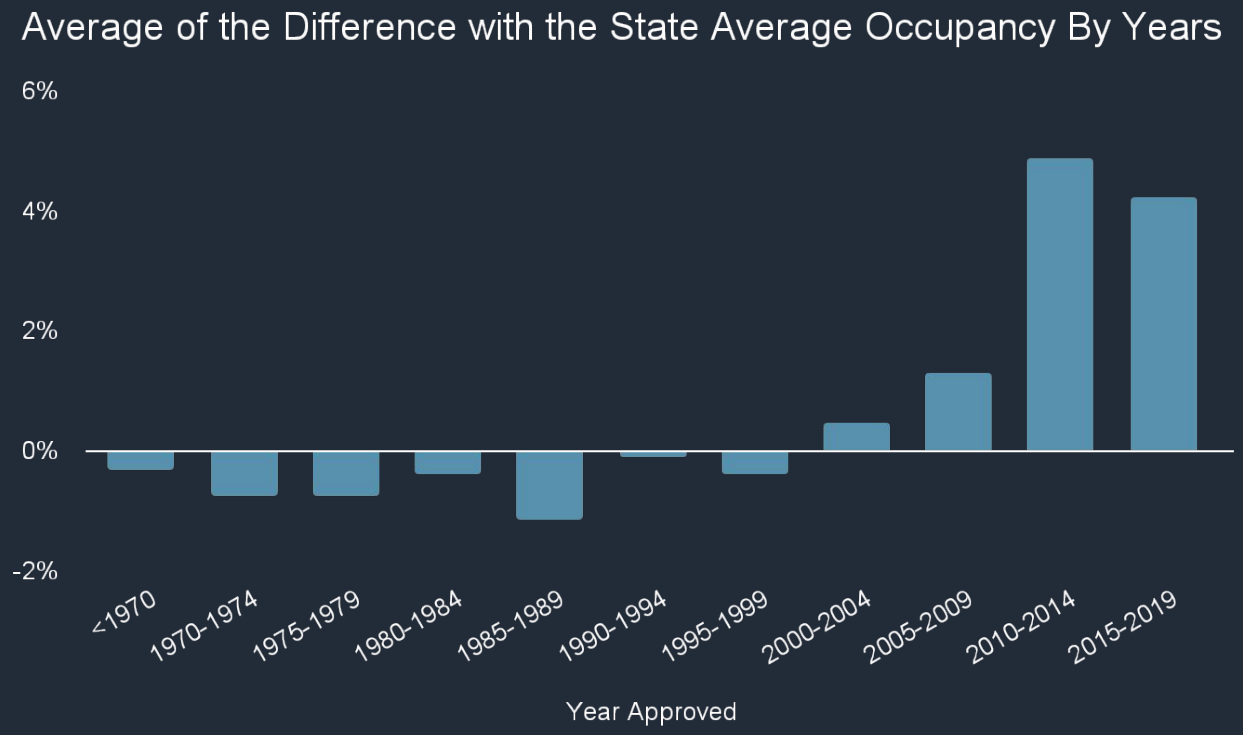
*Data source: Centers for Medicare & Medicaid Services

Preliminary Analysis - Staffing Rating



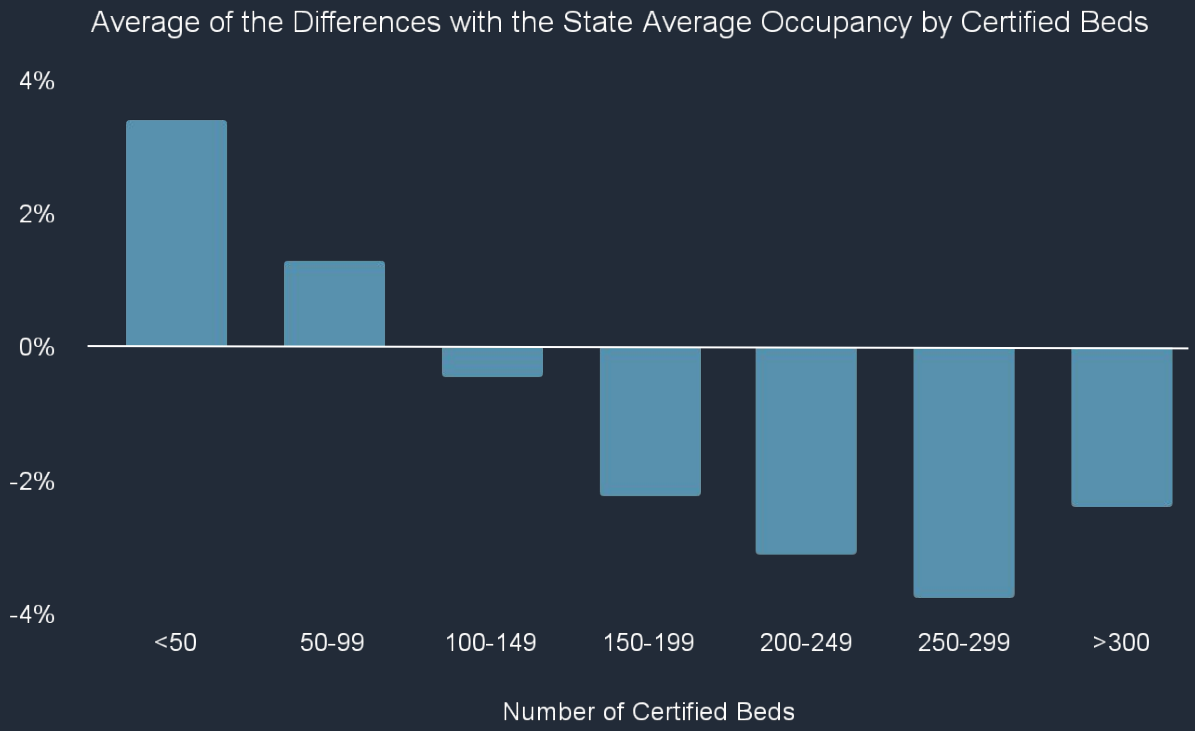
*Data source: Centers for Medicare & Medicaid Services

Preliminary Analysis - Year Approved



*Data source: Centers for Medicare & Medicaid Services

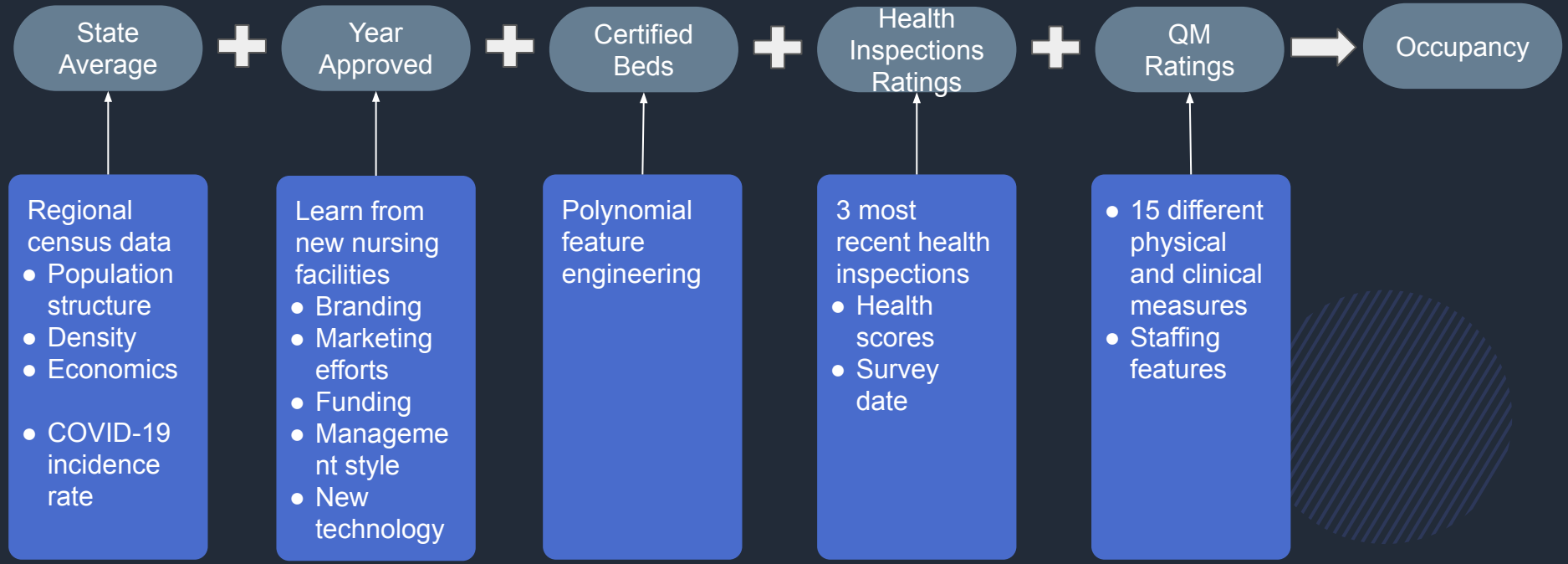
Preliminary Analysis - Certified Beds



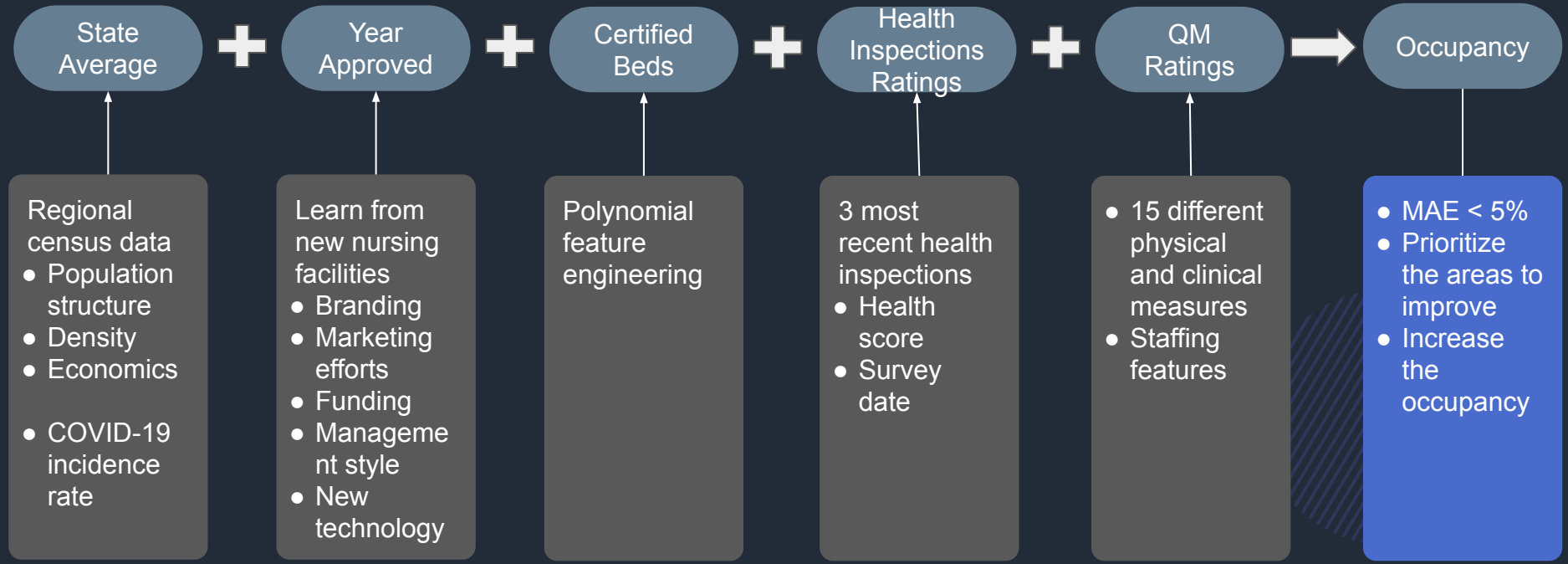
Baseline Model

MAE = 0.105 (10.5%)

Future Work


State Average
Year Approved
Certified Beds
Health Inspections Ratings
QM Ratings
Occupancy
Regional census data
• Population structure
• Density
• Economics
• COVID-19 incidence rate
Learn from new nursing facilities
• Branding
• Marketing efforts
• Funding
• Management style
• New technology
Polynomial feature engineering
3 most recent health inspections
• Health scores
• Survey date
• 15 different physical and clinical measures
• Staffing features

Final Goal - MAE < 5%



Other Opportunities with Data Science

- Fines
- Readmission Rate
- Human Resources

- Many other potentials!

Thank You!

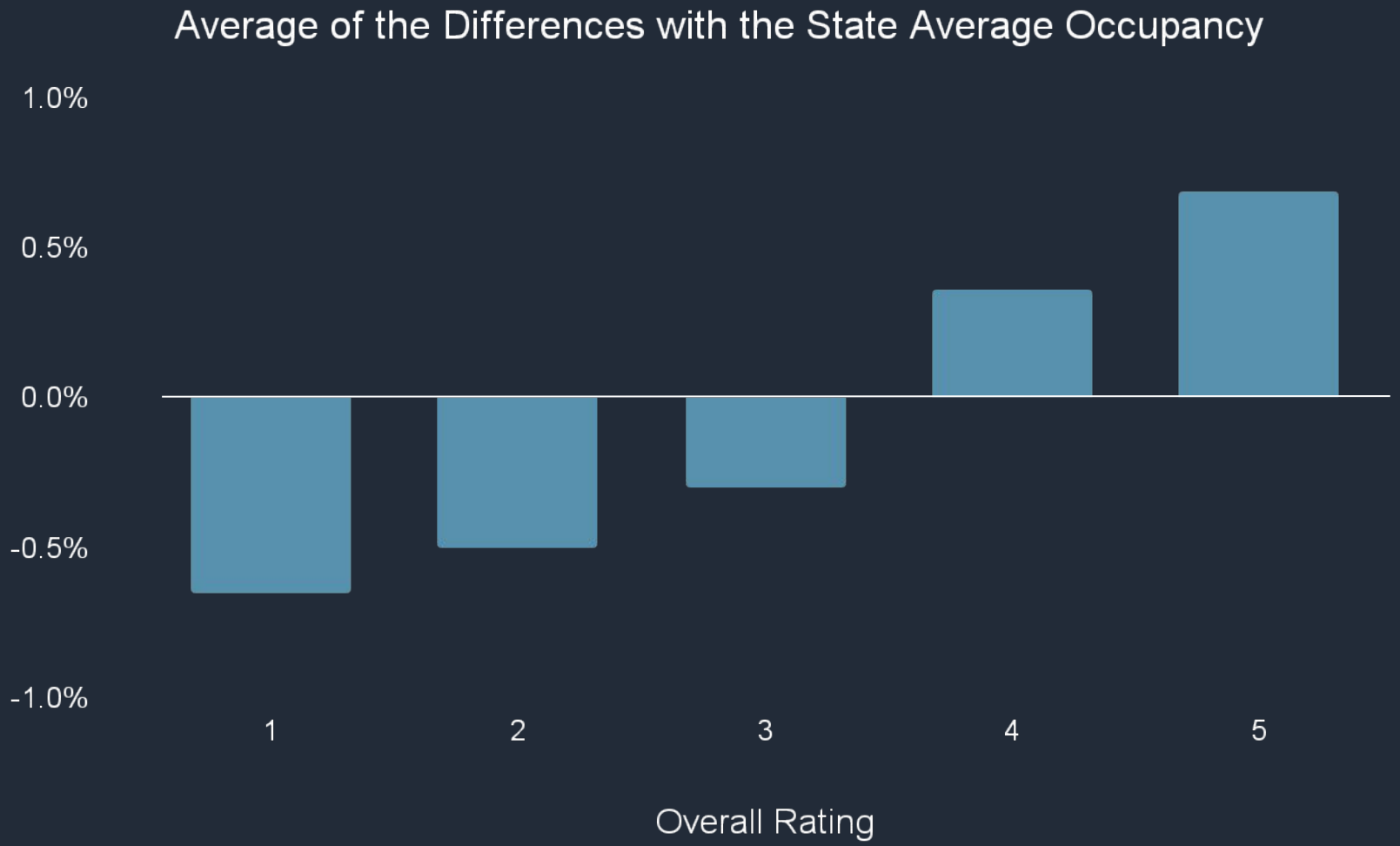
Do you have any questions?

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**.

Appendix

Resources

- Datasets of nursing homes including rehab services on CMS.gov
  - https://data.cms.gov/provider-data/topics/nursing-homes
- Five-Star Quality Rating System
  - https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/CertificationandComplianc/FSQRS
- Tableau Public Dashboard
  - https://public.tableau.com/app/profile/koscew/viz/NursingFacilitiesinU_S_CMSdata/Map



*Data source: Centers for Medicare & Medicaid Services