

Nursing Home Quality Analysis:

Analysis from a Staffing Perspective

by Edwin Ng

Why Nursing Homes?

- Abuse of the elderly is a serious problem observed in institutions such as nursing homes or long-term care facilities around the world.
- According to a “2017 review of 52 studies in 28 countries” it is “estimated that over the past year **1 in 6 people (15.7%)** aged 60 years and older **were subjected to some form of abuse (1)**” (WHO)
- Furthermore, studies indicate that **2 in 3 staff members** of institutions such as nursing homes or long-term care facilities **have admitted to perpetrating some form of abuse** of elderly residents within the last year (WHO)

The Problem:

- It's clear that staff are the ones committing abuse towards their patients, but *why?*
- Working in a retirement home as a caregiver is a **mentally draining** and **physically challenging** job.
- **Residents mistreat** or even **harass/assault** the staff (either as a result of their medical conditions or not).
- Problems such as **understaffing and underpaying**.
- Nursing homes, experienced **turnover of 52%** of nursing staff each year

What is the relationship between Staffing & Care?



How can we use **nursing home staffing rates** to uncover its impact on nursing home quality measures such as **overall care ratings, health inspection ratings, incident reports, and fines/penalties** such that we can make staffing recommendations to local nursing homes within the next year to reduce resident mistreatment and improve staff QOL?



The Data:

- 15640 rows
- 82 columns
- Exclusively US nursing homes.
- From 6 years ago.

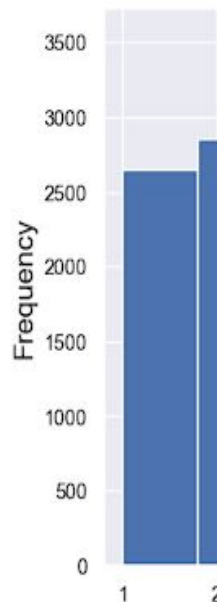
Federal Provider Number	Provider Name	Provider Address	Provider City	Provider State	Provider Zip Code	Provider Phone Number	Provider SSA County Code	Provider County Name	...	Cycle 3 Total Health Score	Total Weighted Health Survey Score
015019	MERRY WOOD LODGE CARE AND REHABILITATION CENTER	P O BOX 130	ELMORE	AL	36025	3345678484	250	Elmore	...	24.0	22.667
015113	RIVER CITY CENTER	1350 FOURTEENTH AVENUE SOUTHEAST	DECATUR	AL	35601	2563556911	510	Morgan	...	40.0	48.667
015112	MAGNOLIA HAVEN HEALTH AND REHABILITATION CENTER	603 WRIGHT STREET	TUSKEGEE	AL	36083	3347274960	430	Macon	...	16.0	15.333
015114	SHADESCREST HEALTH CARE CENTER	331 WEST 25TH STREET	JASPER	AL	35502	2053849086	630	Walker	...	52.0	25.333
015123	WASHINGTON COUNTY NURSING HOME	14600 ST STEPHENS AVENUE	CHATOM	AL	36518	2518476412	640	Washington	...	20.0	13.333

Data Source:

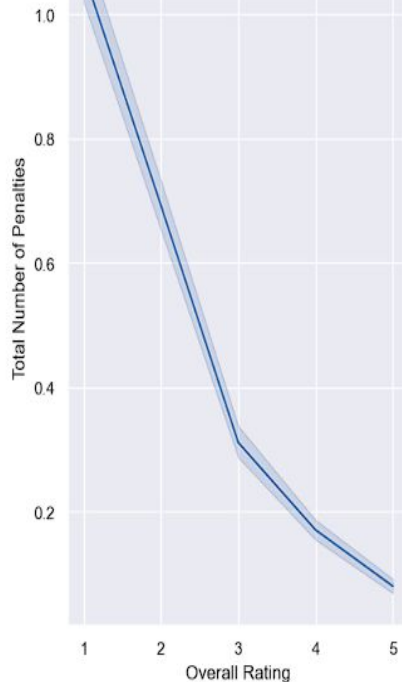
https://www.kaggle.com/datasets/thedevastator/nursing-home-quality-staffing?select=Provider_Info.csv

Data Wrangling:

- Original dataset was 15640 rows x 82 columns.
- Subset dataset to contain only relevant information (staffing hours and responding variables)
- Data validation - ensuring all the same type of data contained within one column.
- No duplicates found.
- Null values in the responding variable were removed.
- 2 null values found in one specific staffing type imputed with average.
- Final dataset was 15183 rows x 41 columns.



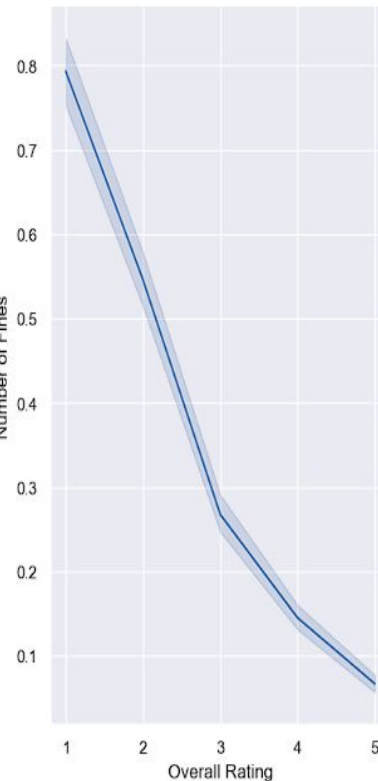
Staffing Ratios



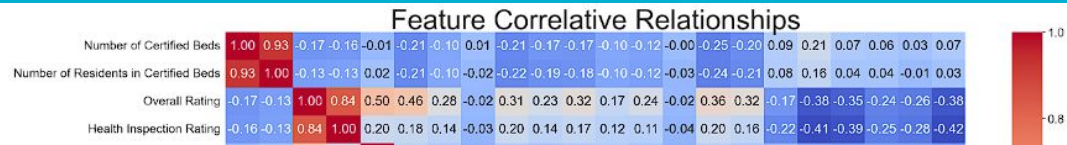
Number of Substantiated Complaints



Number of Fines

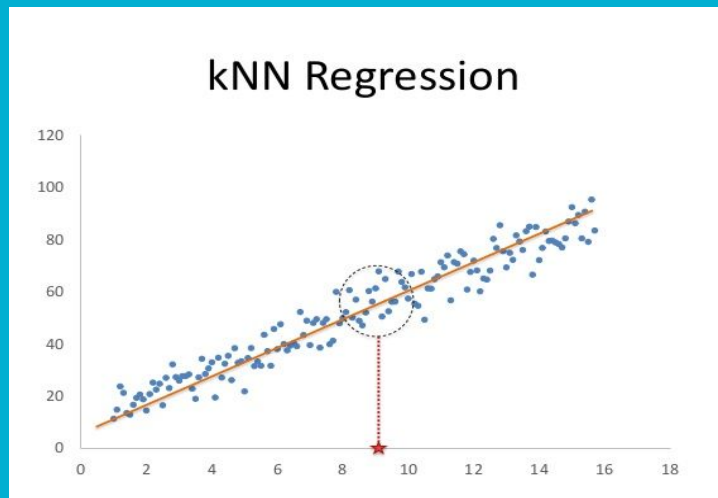
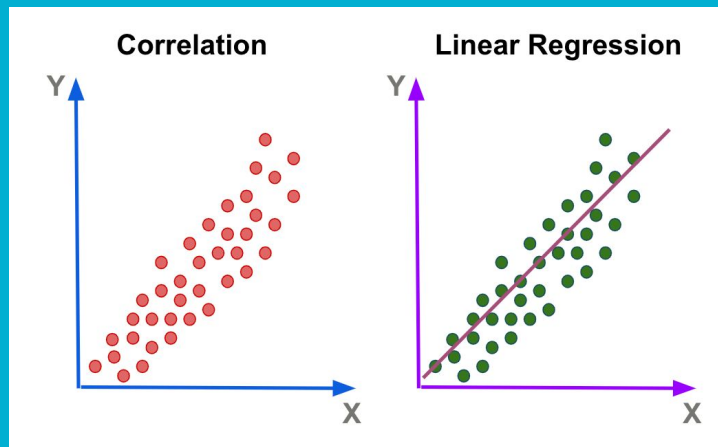
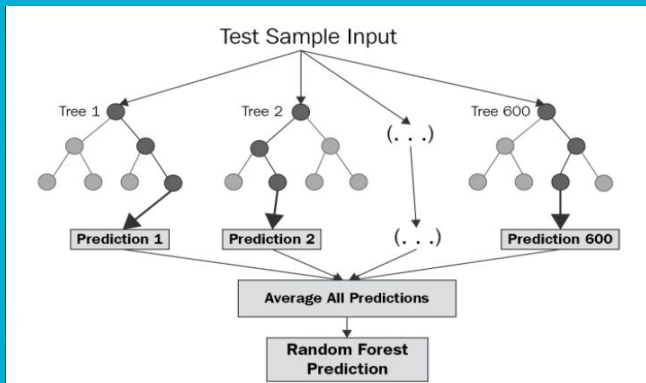


5



Model Selection:

- Linear Regression Model
- K Nearest Neighbors Regression Model
- Random Forest Regression Model



Model Performance:

- Linear Regression:

- The linear regression mean squared error: 0.2126697509640798
- The linear regression model R squared: 0.8948769714186898

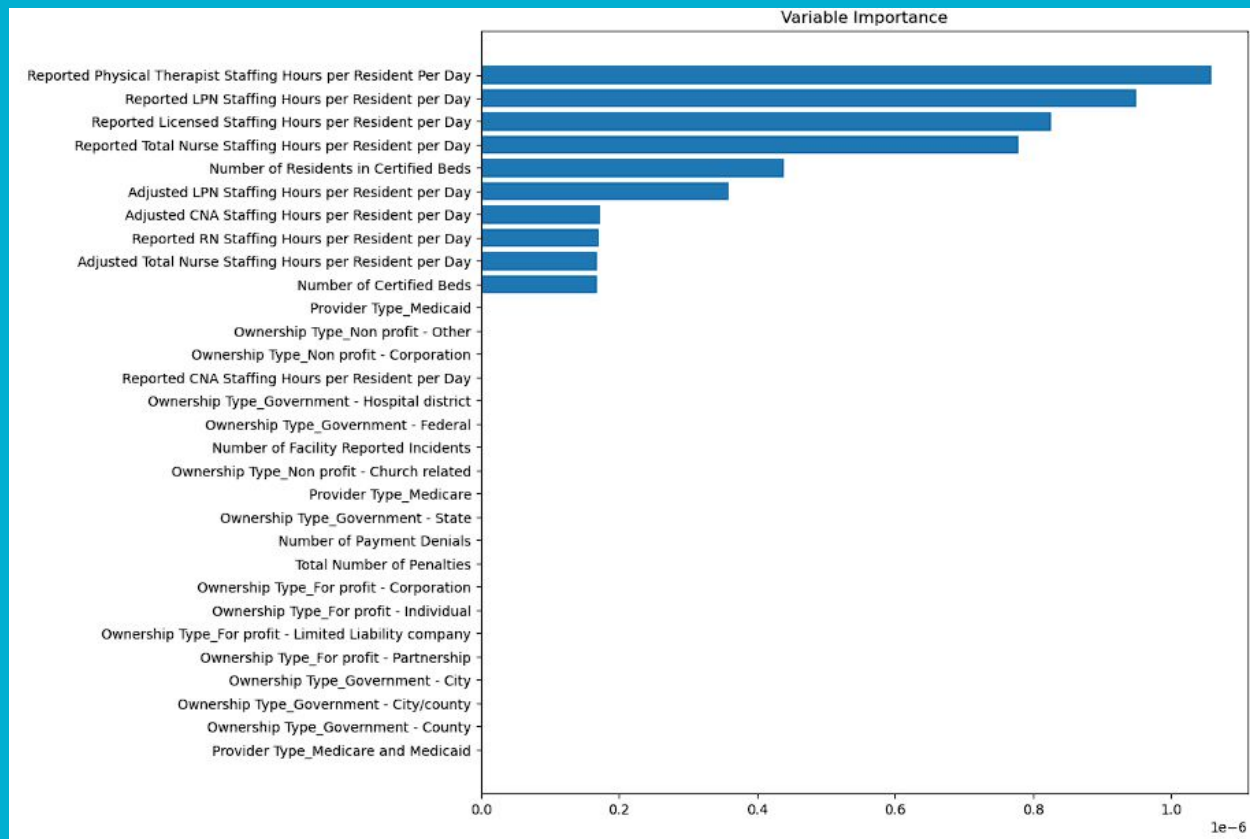
- K-Nearest Neighbors - (172 Neighbors):

- The K Nearest Neighbors model mean squared error: 2.7925584458347052
- The K Nearest Neighbors R squared value: 0.30260125123477116

- Random Forest - (min_samples_leaf=2, min_samples_split=8):

- Random Forest Regressor model mean squared error: 0.0013419260176346691
- Random Forest Regressor R squared value: 0.999336683630529

Feature Importance:



Future Investigations:

- What is the ideal ratio of patients to staff such that **staff are not overworked** and **care quality is optimal**?
- What is the ideal blend of staff at a nursing home? What **ratio of each profession** within a nursing home leads to the **best quality**?
- Given a **realistic staffing budget** how can we create the best team?

Thank You!

Special Thanks to:

- Kaggle for creating readily accessible free datasets
- David Lara-Arango for mentoring throughout this project
- Data.World for hosting the original source dataset