

Practicum

Correlation Analysis on Nursing Turnover¹

Practicum

Correlation Analysis on Nursing Turnover

Nick Molliconi

nmolliconi@yahoo.com

Regis University

MSDS 692 Data Science Practicum

Dr. Ksenia Polson

March 5, 2023

Practicum

Correlation Analysis on Nursing Turnover2

INTRODUCTION:

Since the outbreak of the Covid-19 pandemic, there has been an ever-increasing shortage of nurses with many leaving the profession (Gaines, 2022). Reasoning behind the significant departure, and shortage, are due to many factors. In the article, “What’s really behind the nursing shortage? 1,500 Nurses Share Their Story,” states that many factors, such as inadequate staffing ratios, pay inequities and lack of time off (Gaians, 2022) are causing nurses to leave their profession. This project will look at factors, such as age and where nurses work in an acute care setting to determine if there is a strong correlation between these factors and leaving the profession. Specifically, those nurses who are older than forty-five (45) and nurses who work in critical care units, such as the Emergency Department (ED) and Intensive Care Unit (ICU) have a higher likelihood of leaving the profession compared to those Nurses who are not in those two groups.

DATASET:

The data was derived from a national survey created by Dr. Deb Roybal, PhD RN MS, Associate Professor at the Regis University Loretto Heights School of Nursing and Ksenia Polson Ph.D. Assistant Professor at the Regis University Anderson College of Business and Computing. The survey tool consists of forty-two (42) questions aimed at gathering data to look into why nurses are leaving the bedside. An issue for this project is a lower response rate from the survey, and of those who did respond, did they respond to all the questions within the survey.

ANALYSIS:

The first hypothesis is that older aged nurses, ages from forty-five and older (45+), have a higher probability to leave the profession more than their younger peers, ages forty-five (45) and

Practicum

Correlation Analysis on Nursing Turnover³

below. Additionally, the second hypothesis is that those nurses who work in critical care departments in acute care settings, such as the Emergency Department and Intensive Care Units (ICU) have a higher probability of leaving than those nurses who are not in critical care departments. Additional hypotheses will look at both length of having a license of five (5) years or more and current salary with the higher probability of leaving.

As stated above, the features from the survey will include age groupings, departments within an acute Care setting and the decision to leave the profession on the survey tool. Additional features will look at other care settings and the decision to leave the profession and demographic data and visualizations. The analysis will include comparative analysis utilizing a Chi-squared contingency analysis around the following alternative hypotheses:

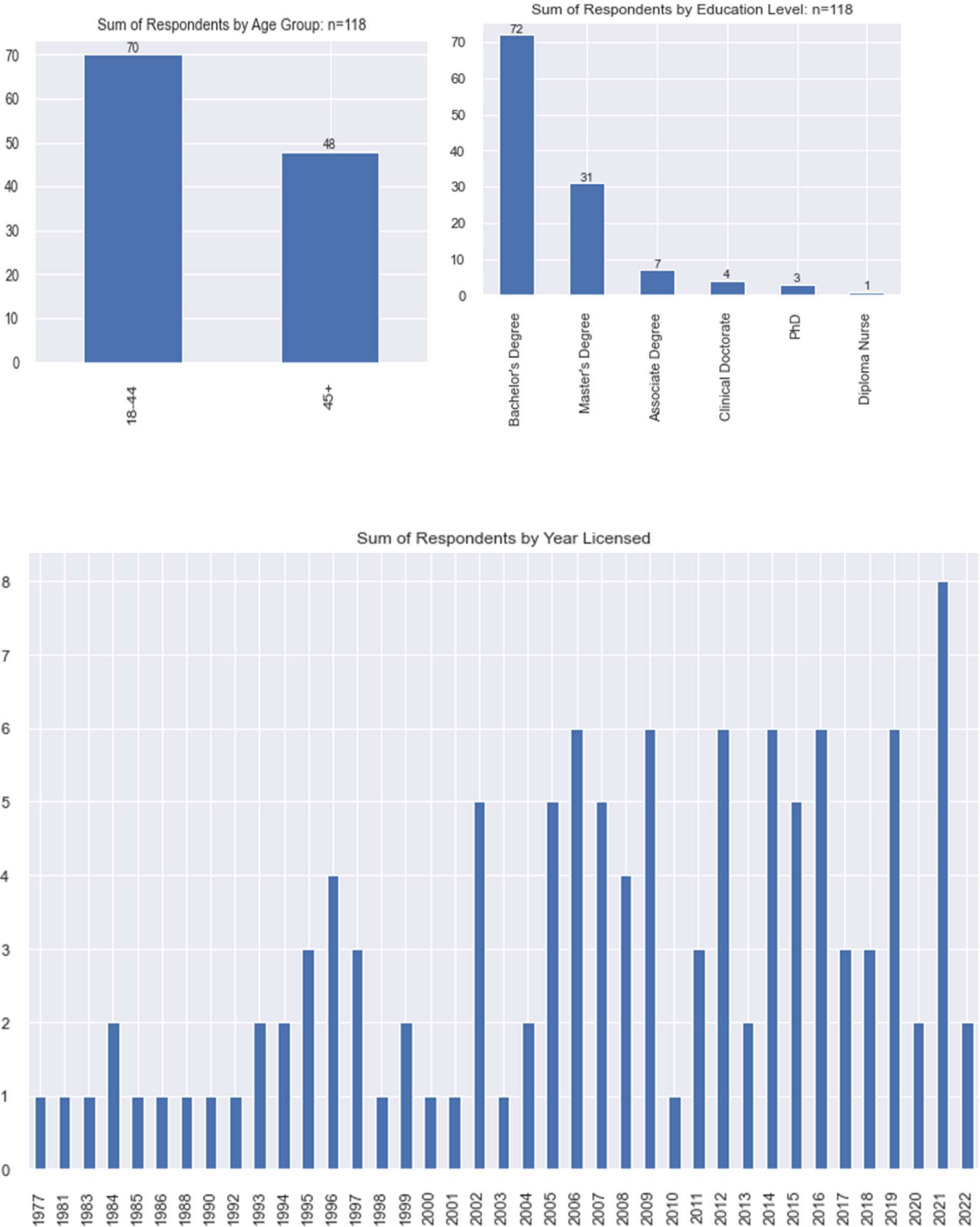
1. Variables of age and leaving within six (6) months will have a high correlation.
2. Acute care hospital departments and leaving within six months will have a high correlation.
3. Years of being licensed and leaving within six (6) months will have a high correlation.
4. Current salary of respondents and leaving within six (6) months will have a high correlation.
5. Variables of age and changed jobs within last six (6) months will have a high correlation.

Additional analysis included the Spearman Correlation to test to dependence between categorical columns.

Practicum

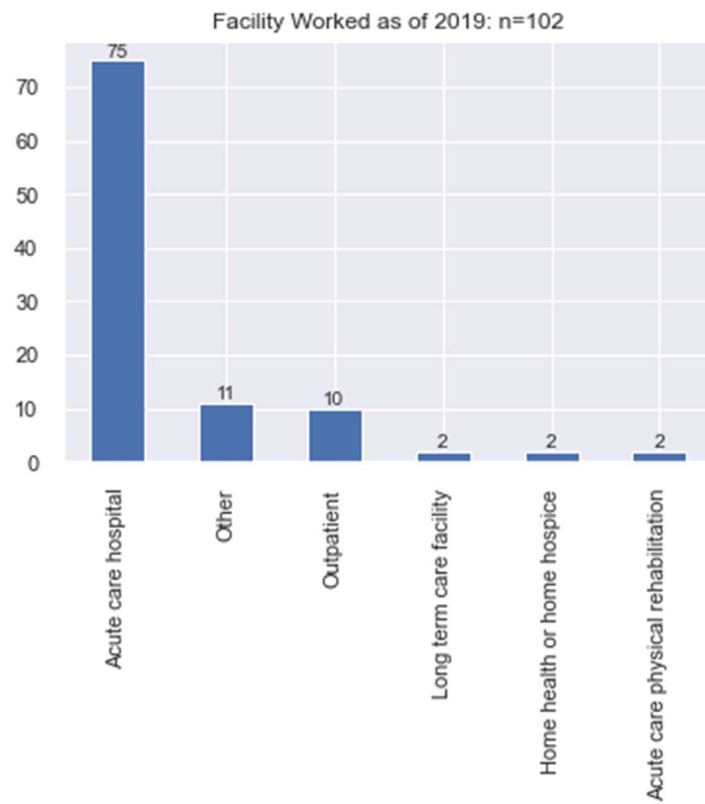
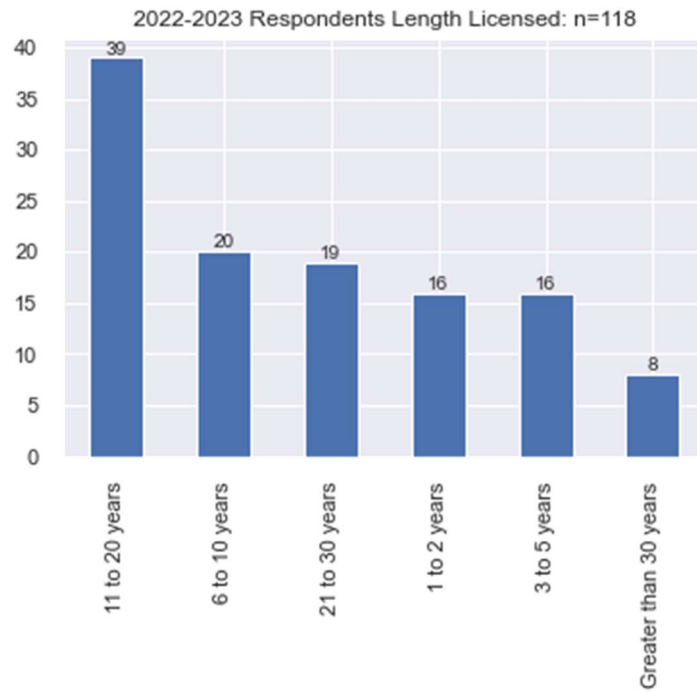
Correlation Analysis on Nursing Turnover4

EDA:



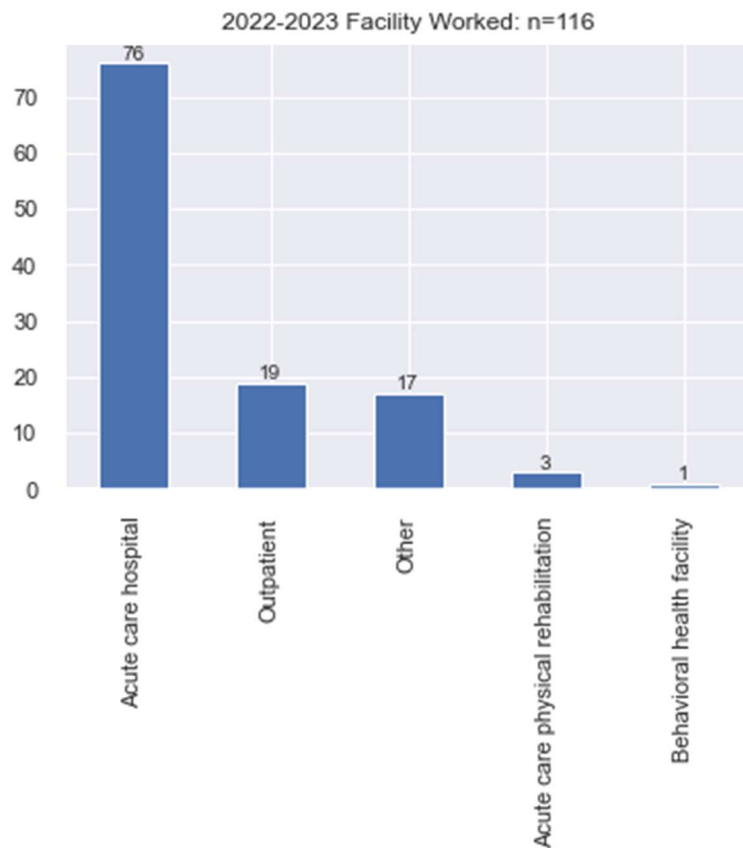
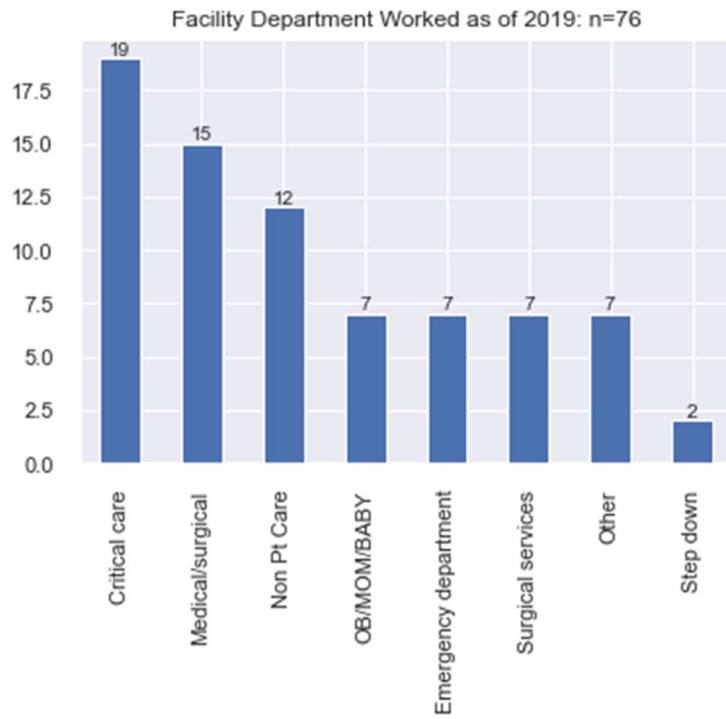
Practicum

Correlation Analysis on Nursing Turnover5



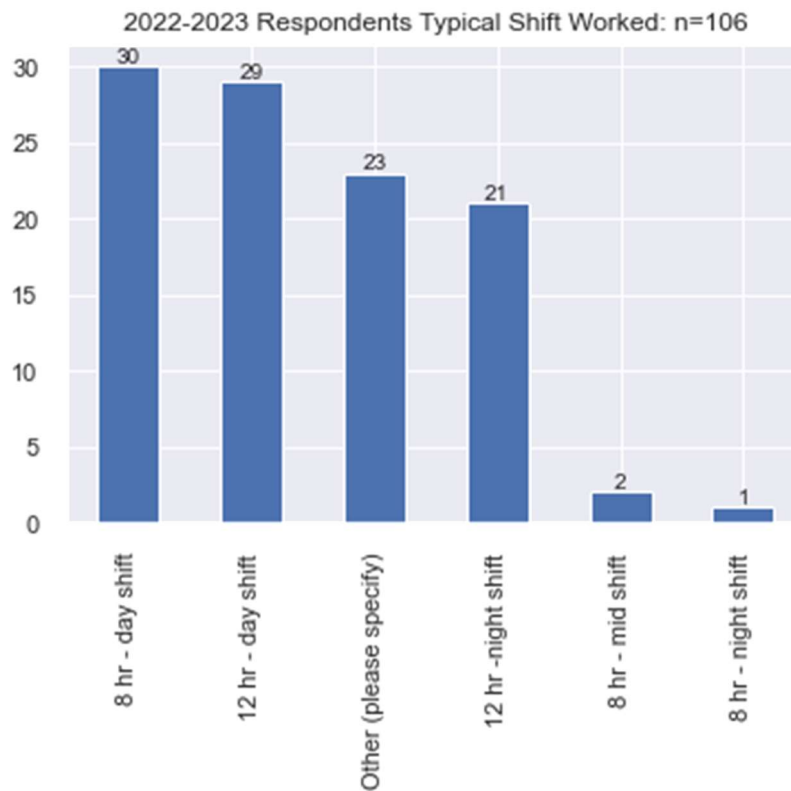
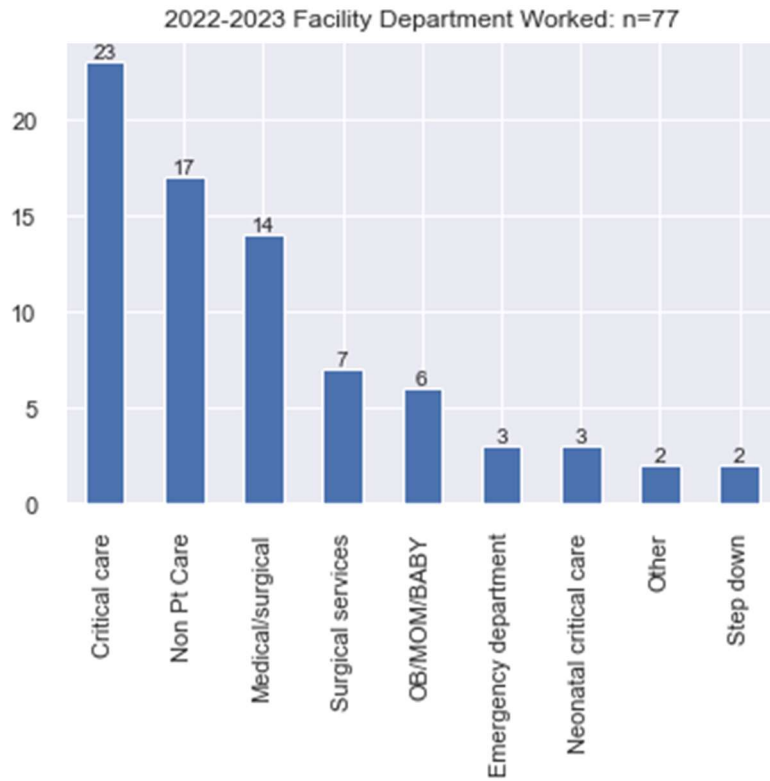
Practicum

Correlation Analysis on Nursing Turnover6



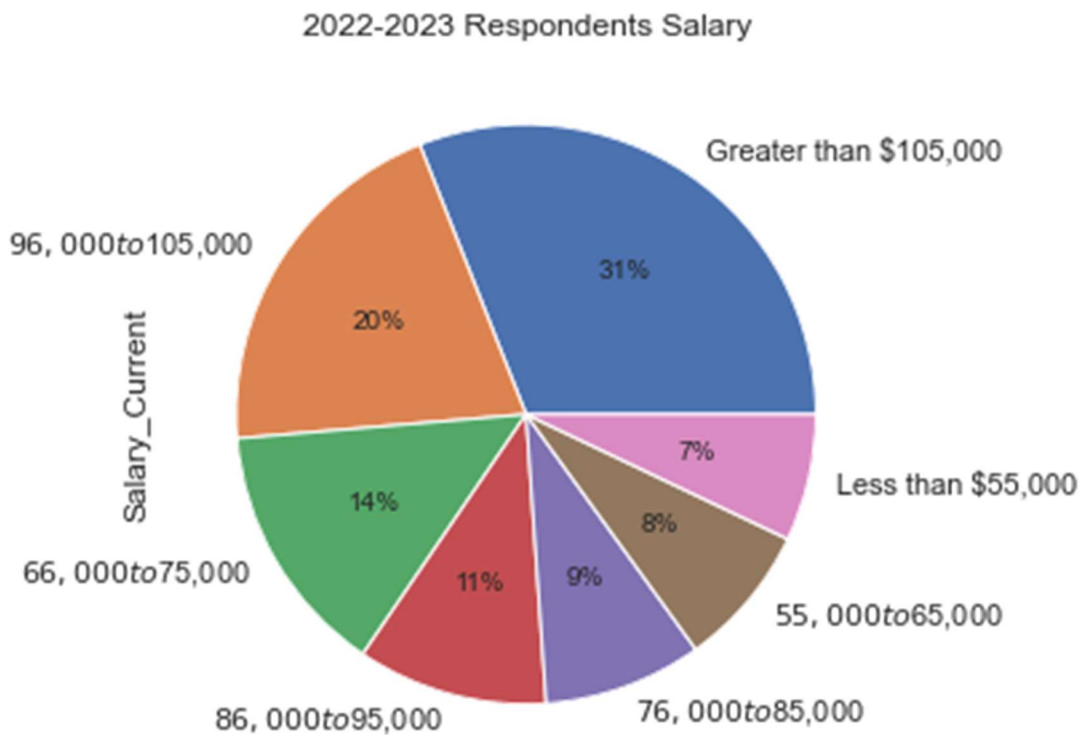
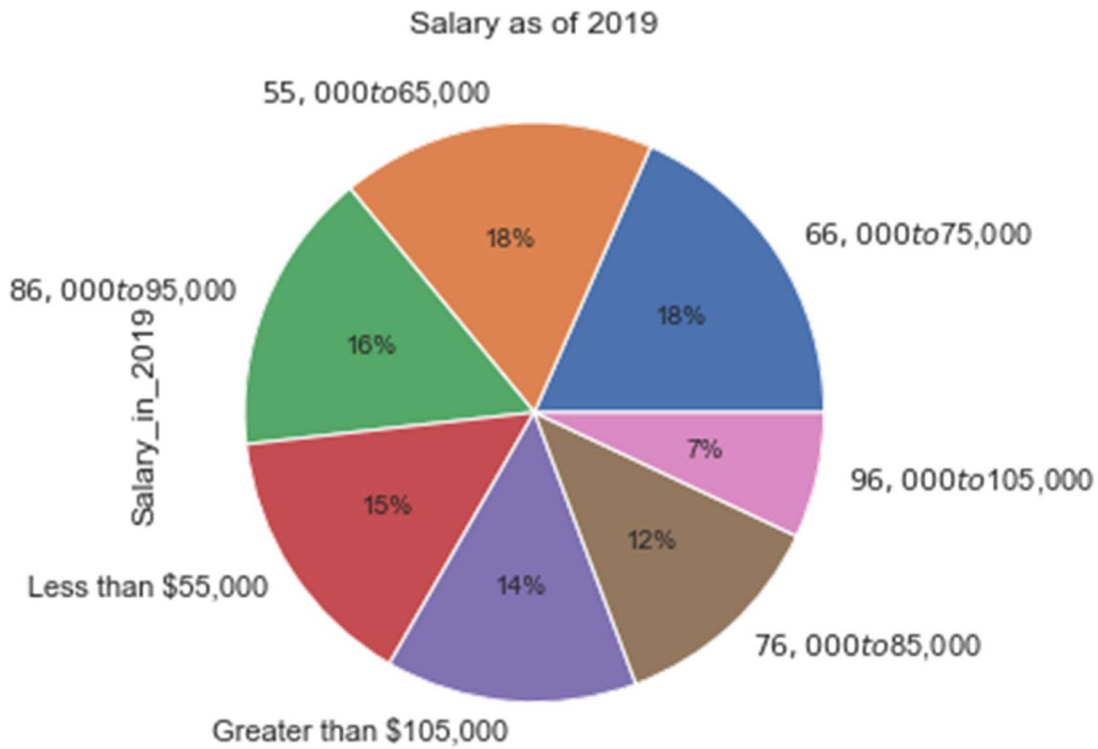
Practicum

Correlation Analysis on Nursing Turnover7



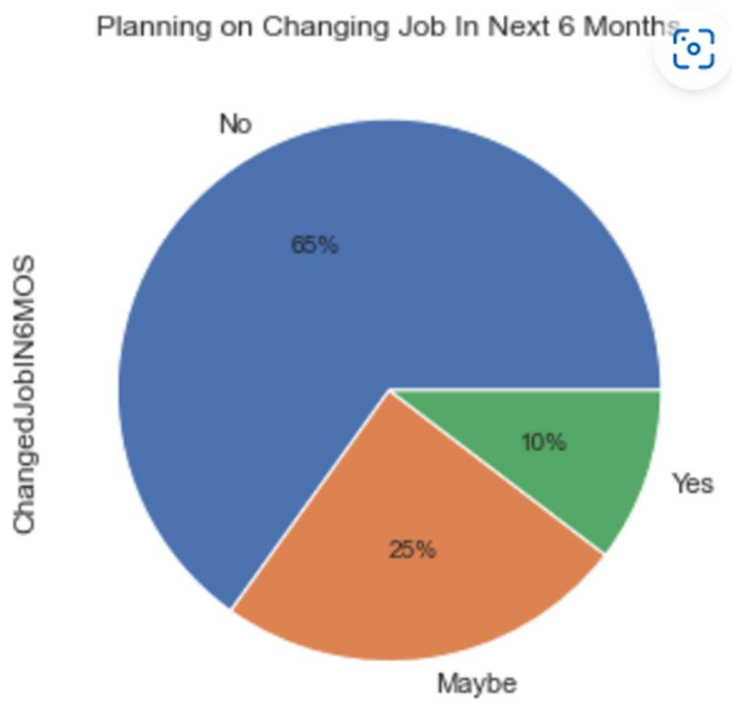
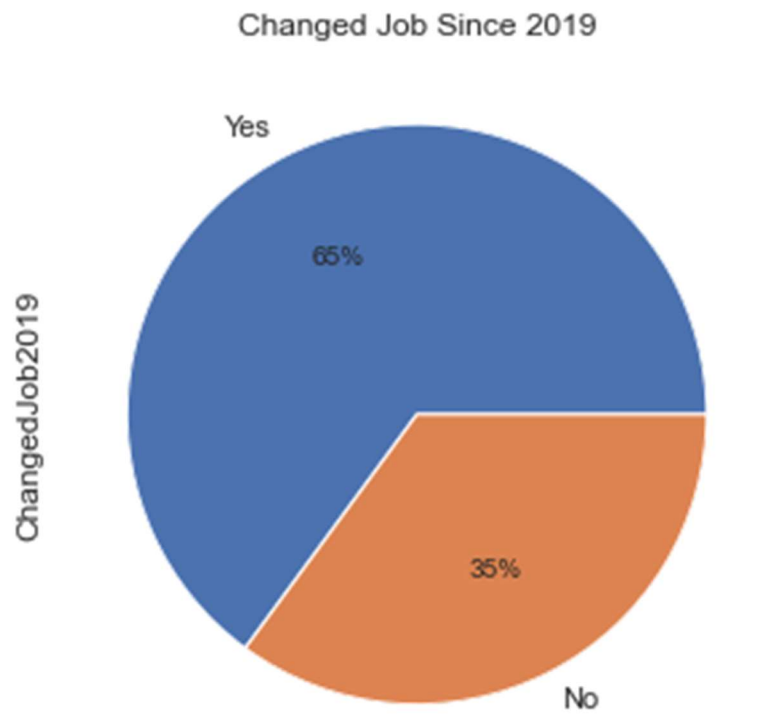
Practicum

Correlation Analysis on Nursing Turnover8



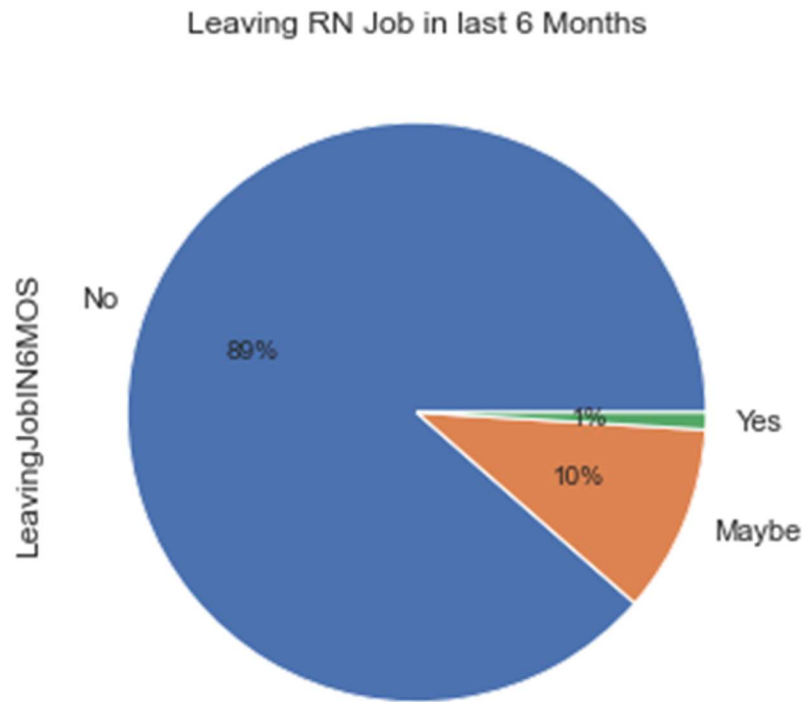
Practicum

Correlation Analysis on Nursing Turnover9



Practicum

Correlation Analysis on Nursing Turnover10



Practicum

Correlation Analysis on Nursing Turnover¹¹

ANALYSIS:

Below is a summary table that provides the results for all of the Chi-square and Sherman's Rank Correlation tests:

	Leaving Position in Six (6) Months				a = 0.05	NULL HYPOTHESIS
	X ²	p=	r	p=		
Age Group	0.83	0.66	-0.008	0.93	p > a	FAIL TO REJECT
Current Acute Care Department	15.1	0.37	-0.005	0.97	p > a	FAIL TO REJECT
Years of Being Licensed	9.65	0.29	0.105	0.29	p > a	FAIL TO REJECT
Current Salary	8.98	0.7	-0.14	0.15	p > a	FAIL TO REJECT

	Changed Jobs in Past Six (6) Months				a = 0.05	NULL HYPOTHESIS
	X ²	p=	r	p=		
Age Group	12.9	0.23	-0.091	0.35	p > a	FAIL TO REJECT

H₀: Variables of age and leaving within six (6) months will not have a high correlation.

H₁: Variables of age and leaving within six (6) months will have a high correlation.

```
c, p, dof, expected = chi2_contingency(contingency_pct)
p
```

```
0.6614954723209638
```

```
print(c)
```

```
0.8265042800289512
```

```
coef, p = spearmanr(RNAGE5['AgeGroup'], RNAGE5['LeavingJobIN6MOS'])
seed(1)
print('Spearman's correlation coefficient: %.3f' % coef)
alpha = 0.05
if p > alpha:
    print('Samples are uncorrelated (fail to reject H0) p=%.3f' % p)
else:
    print('Samples are correlated (reject H0) p=%.3f' % p)
```

```
Spearman's correlation coefficient: -0.008
Samples are uncorrelated (fail to reject H0) p=0.935
```

Practicum

Correlation Analysis on Nursing Turnover¹²

Both the Chi-squared and Spearman correlation tests indicate that both the Age Group and leaving your job in six (6) months columns are independent of each other, thus failing to reject the Null hypothesis.

$$\chi^2(1, N=105) .83, p = .66.$$

$$r(103) = -.008, p=.935$$

H0: Acute care hospital departments and leaving within six months will not have a high correlation.

H1: Acute care hospital departments and leaving within six months will have a high correlation.

```
c, p, dof, expected = chi2_contingency(contingency_pct)
p
```

```
0.3694925229113806
```

```
print(c)
```

```
15.127534562211983
```

```
coef, p = spearmanr(RNDEPT['Current_Facility_Department_Worked'], RNDEPT['LeavingJobIN6MOS'])
seed(1)
print('Spearman correlation coefficient: %.3f' % coef)
alpha = 0.05
if p > alpha:
    print('Samples are uncorrelated (fail to reject H0) p=%.3f' % p)
else:
    print('Samples are correlated (reject H0) p=%.3f' % p)
```

```
Spearman correlation coefficient: -0.005
Samples are uncorrelated (fail to reject H0) p=0.968
```

Both the Chi-squared and Spearman correlation tests indicate that both the Acute care hospital departments and leaving your job in six (6) months columns are independent of each other, thus failing to reject the Null hypothesis.

$$\chi^2(1, N=69)15.1, p = .37.$$

Practicum

Correlation Analysis on Nursing Turnover¹³

$r(67) = -.005, p = .968$

H0: Years of being licensed and leaving within six (6) months will not have a high correlation.

H1: Years of being licensed and leaving within six (6) months will have a high correlation.

```
c, p, dof, expected = chi2_contingency(contingency_pct)
p
```

```
0.2906215608562678
```

```
print(c)
```

```
9.647789229514803
```

```
coef, p = spearmanr(RNYL['Current_Level_Length_Licensed'], RNYL['LeavingJobIN6MOS'])
seed(1)
print('Spearman correlation coefficient: %.3f' % coef)
alpha = 0.05
if p > alpha:
    print('Samples are uncorrelated (fail to reject H0) p=%.3f' % p)
else:
    print('Samples are correlated (reject H0) p=%.3f' % p)
```

```
Spearman correlation coefficient: 0.105
Samples are uncorrelated (fail to reject H0) p=0.288
```

Both the Chi-squared and Spearman correlation tests indicate that both the number of years licensed and leaving your job in six (6) months columns are independent of each other, thus failing to reject the Null hypothesis.

$\chi^2(1, N=105) 9.64, p = .290$.

$r(103) = .105, p = .288$

H0: Current salary of respondents and leaving within six (6) months will not have a high correlation.

H1: Current salary of respondents and leaving within six (6) months will have a high correlation.

Practicum

Correlation Analysis on Nursing Turnover¹⁴

```
c, p, dof, expected = chi2_contingency(contingency_pct)
p
```

```
0.7043332507178435
```

```
print(c)
```

```
8.983568624383055
```

```
coef, p = spearmanr(RNSAL['Salary_Current'], RNSAL['LeavingJobIN6MOS'])
seed(1)
print('Spearman correlation coefficient: %.3f' % coef)
alpha = 0.05
if p > alpha:
    print('Samples are uncorrelated (fail to reject H0) p=%.3f' % p)
else:
    print('Samples are correlated (reject H0) p=%.3f' % p)
```

```
Spearman correlation coefficient: -0.142
Samples are uncorrelated (fail to reject H0) p=0.151
```

Both the Chi-squared and Spearman correlation tests indicate that both the current salary of respondents and leaving your job in six (6) months columns are independent of each other, thus failing to reject the Null hypothesis.

$\chi^2(1, N=104) 8.98, p = .704$

$r(102) = -0.14, p = .151$

H0: Variables of age and changed jobs within last six (6) months will not have a high correlation.

H1: Variables of age and changed jobs within last six (6) months will have a high correlation.

Practicum

Correlation Analysis on Nursing Turnover15

```
c, p, dof, expected = chi2_contingency(contingency_pct)
p
```

```
0.23137843776268407
```

```
print(c)
```

```
12.863948215308664
```

```
coef, p = spearmanr(RNNEW['AgeGroup'], RNNEW['ChangedJobIN6MOS'])
seed(1)
print('Spearman correlation coefficient: %.3f' % coef)
alpha = 0.05
if p > alpha:
    print('Samples are uncorrelated (fail to reject H0) p=%.3f' % p)
else:
    print('Samples are correlated (reject H0) p=%.3f' % p)
```

```
Spearman correlation coefficient: -0.091
```

```
Samples are uncorrelated (fail to reject H0) p=0.352
```

Both the Chi-squared and Spearman correlation tests indicate that both the age group and changed your job in six (6) months columns are independent of each other, thus failing to reject the Null hypothesis.

$\chi^2(1, N=106)$ 12.9, $p = .231$

$r(104) = -0.09$, $p = .352$

Correlation Analysis on Nursing Turnover16

SENTIMENT ANALYSIS:

Below are three-word clouds that reflect the sentiment responses for questions relating to what respondents least like about being a registered nurse (RN), what they find challenging about being a RN, and what they find rewarding on being a RN.

Least Like About Being a Registered Nurse (RN):



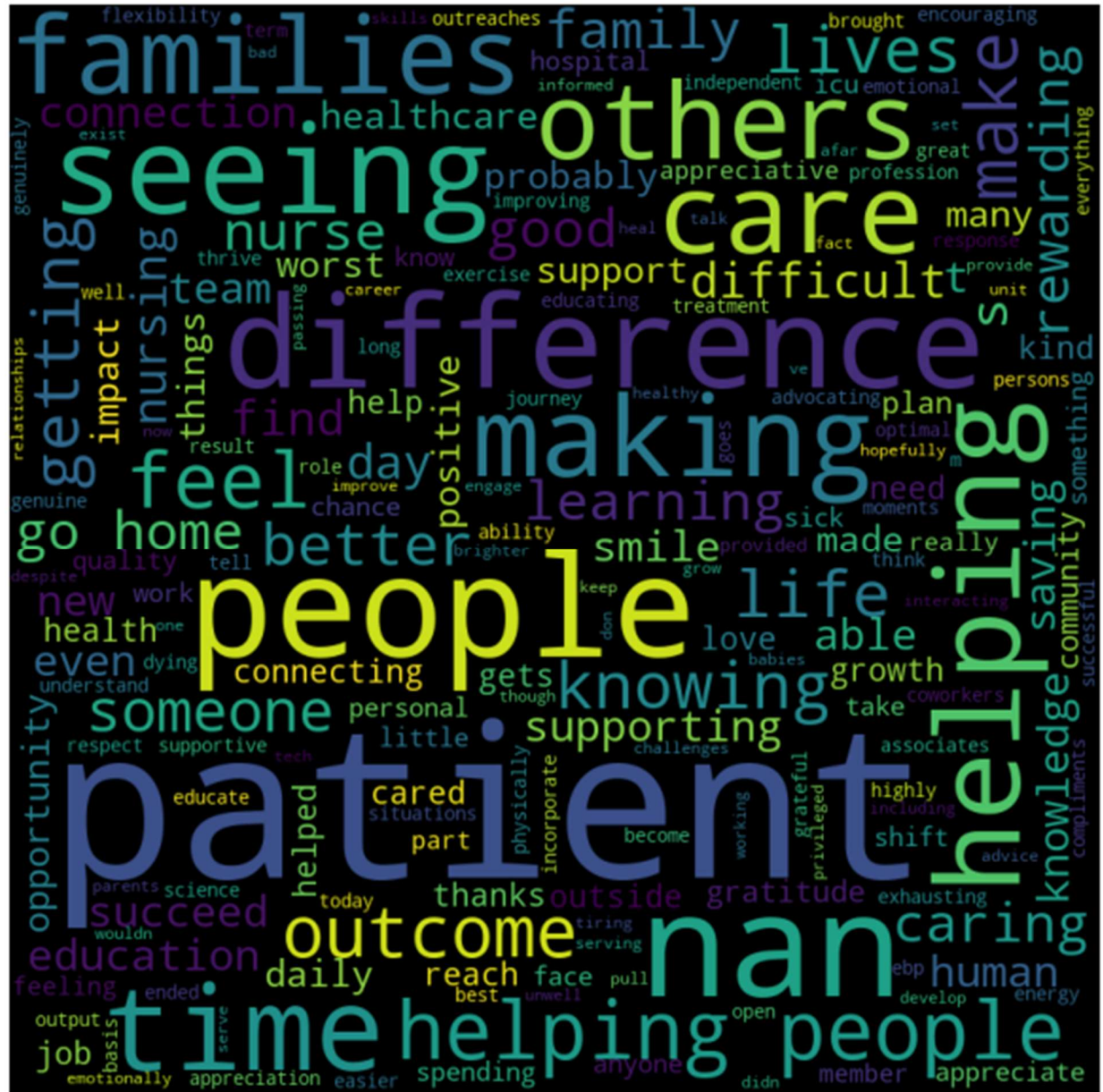
Correlation Analysis on Nursing Turnover¹⁷

Challenging About Being a Registered Nurse (RN):



Correlation Analysis on Nursing Turnover¹⁸

Rewarding About Being a Registered Nurse (RN):



Practicum

Correlation Analysis on Nursing Turnover¹⁹

CONCLUSION:

The results from the two correlation tests indicate that none of the variables are dependent on either respondent leaving their respective jobs in the next six months nor planning on changing jobs in the next six months. Each of the hypotheses stated that there would be a level of dependance, or correlation, and none proved true. The variables of age group, facility department and length of having a license used the entire population of respondents and were not broken down into sub groups. The initial design was to use subgroups for each of these variables such as age 45+, critical care departments and nurses licensed five (5) years or less, but due to a small response rate, the total responses were used for those variables.

The sentiment analysis word clouds provided some context to the overall study. The three-word cloud visualizations represented questions around what nurses like about their jobs, what they don't like and what are their challenges. For each of them, the word Patient was the dominant word, but not necessarily meaning that nurses felt challenged, or did not like patients, it was potentially just a common word to their respective responses to the questions asked. However, it may perhaps lead to more analysis around how nurses value relationships more than any another of the variables that were asked in the survey, such as salary, length of being licensed or where they work.

Practicum

Correlation Analysis on Nursing Turnover20

REFERENCES;

Brownlee, Jason. (2019, August 8). How to Calculate Nonparametric Rank Correlation in Python. [www.machinelearningmaster.com](https://machinelearningmastery.com/how-to-calculate-nonparametric-rank-correlation-in-python/#:~:text=Spearman%E2%80%99s%20rank%20correlation%20can%20be%20calculated%20in%20Python,p-value%20for%20interpreting%20the%20significance%20of%20the%20coefficient.). Retrieved on February 20, 2023. <https://machinelearningmastery.com/how-to-calculate-nonparametric-rank-correlation-in-python/#:~:text=Spearman%E2%80%99s%20rank%20correlation%20can%20be%20calculated%20in%20Python,p-value%20for%20interpreting%20the%20significance%20of%20the%20coefficient.>

Data Draft. (2020, December 16). Python for Data Analysis: Chi-Squared Tests. [www.bing.com](https://www.bing.com/videos/search?q=how+do+I+list+out+the+chi+squared+value+in+python&docid=608017522627908236&mid=4C4F1ED9A9167C8C68A54C4F1ED9A9167C8C68A5&view=detail&FORM=VIRE). Retrieved on February 15, 2023 from <https://www.bing.com/videos/search?q=how+do+I+list+out+the+chi+squared+value+in+python&docid=608017522627908236&mid=4C4F1ED9A9167C8C68A54C4F1ED9A9167C8C68A5&view=detail&FORM=VIRE>.

Gaines, Kathleen. (2022, October 10. What's Really Behind the Nursing Shortage? 1,500 Nurses Share Their Story. [www.nurse.org](https://nurse.org/articles/nursing-shortage-study/). Retrieved on January 12, 2022, from <https://nurse.org/articles/nursing-shortage-study/>.

Pipis, George. (2020, September 1). How to Run Chi-Square Test in Python. [www.predictivehacks.com](https://predictivehacks.com/how-to-run-chi-square-test-in-python/). Retrieved on February 17, 2023. <https://predictivehacks.com/how-to-run-chi-square-test-in-python/>.