

# Analyzing US Healthcare Providers

## Abstract

We notice that there are important gender differences between male and female healthcare providers. We also find that there is variation among states in the availability of Magnetic Resonance Imaging (MRI) services with Florida and Connecticut having the highest and lowest MRI densities respectively.

## Introduction

In this project, we analyze healthcare providers available in the US. We use the National Plan and Provider Enumeration System (NPPES) database, a publicly available data source which reports the status of the healthcare providers across the across. The federal government mandates all healthcare providers to update important information regarding their services. These data points include the type of provider, location, gender (for individual providers), and the types of services provided to patients. This database is very extensive, and analysis of this data would allow us to get a better understanding of the characteristics of healthcare providers in the US. We start out with a brief summary of our results. This is followed by a comprehensive analysis of the results, and a set of recommendations for state and federal governments. We complete our analysis with a conclusion and a list of the references used in preparing this report.

## Results and Discussions

### Statistical test for sole proprietor- Assigned states only

	Female	Male	<i>Total</i>
<b>Sole proprietor</b>	81915	53867	<b>135782</b>
<b>% within</b>	16.6%	10.9%	<b>27.5%</b>
<b>Partnership</b>	218305	139936	<b>358241</b>
<b>% within</b>	44.2%	28.3%	<b>72.5%</b>
<b><i>Total</i></b>	<b>300220</b>	<b>193803</b>	<b>494023</b>
<b>% total</b>	<b>60.8%</b>	<b>39.2%</b>	<b>100%</b>

Fisher's test result:

#### Fisher's Exact Test for Count Data

```
data: df2
p-value = 8.996e-05
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.9624002 0.9873441
sample estimates:
odds ratio
 0.9747673
```

#### Analysis:

After analyzing the gender and sole proprietor data of our assigned states into the 2x2 cross-table above, we found that there are more females in both sole proprietor and partnership categories since the population contains more females. Intuitively, we didn't see much gender preference difference from the 2x2 cross table since the number of females in both sole proprietor and partnership are larger than the number of males. However, Fisher's Exact Test for this table shows that gender preference difference is actually significant at the 5% level since the p-value is less than 0.05. Moreover, the odds ratio is less than 1, which means females are less likely than males to start a sole practice office.

#### Statistical test on risk and reward- Assigned states only

	Female	Male	<i>Total</i>
<b>High risk</b>	1557	7553	<b>9110</b>
<b>% within</b>	6.3%	30.5%	<b>36.8%</b>
<b>Low risk</b>	9735	5912	<b>15647</b>
<b>% within</b>	39.3%	23.9%	<b>54.4%</b>
<b><i>Total</i></b>	<b>11292</b>	<b>13465</b>	<b>24757</b>
<b>% total</b>	<b>45.6%</b>	<b>54.4%</b>	<b>100.0%</b>

#### Fisher's test result:

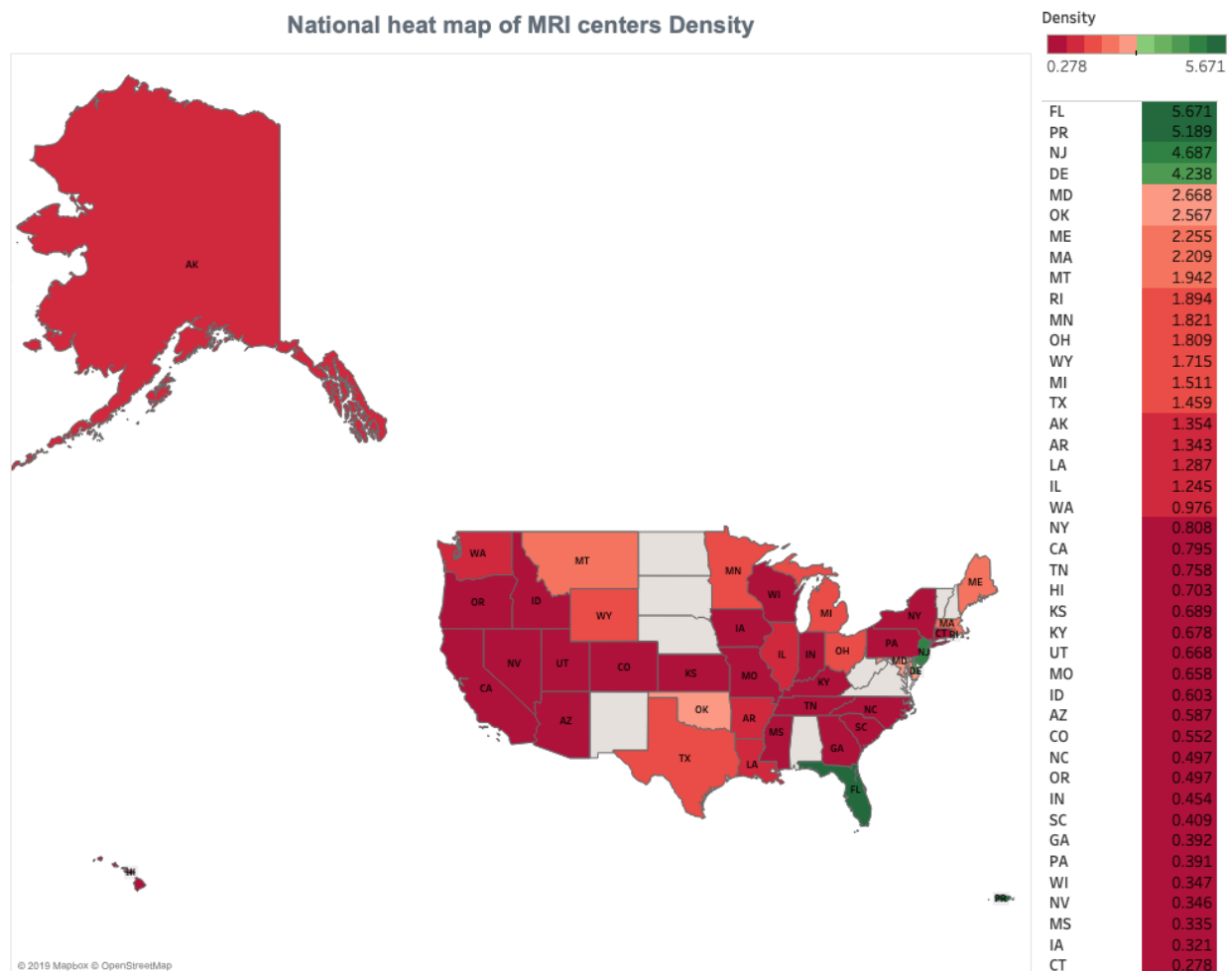
#### Fisher's Exact Test for Count Data

```
data: df4
p-value < 2.2e-16
alternative hypothesis: true odds ratio is less than 1
95 percent confidence interval:
 0.0000000 0.1321261
sample estimates:
odds ratio
 0.1252007
```

### Analysis:

When we look into the table shown above, there is one thing that cannot be ignored. In general, females are more inclined to participate in the practices associated with lower risks. On the contrary, males demonstrate a slightly higher tendency to engage in practices associated with higher risks, which, in return, have higher rewards. To further examine the hypothesis that male doctors are more likely than their female peers to choose the practices that are associated with a higher risk for a higher reward, we ran Fisher's Exact Test on gender and practices associated with higher risks and lower risks. The outcome turns out that the odds ratio in this case is 0.1252, suggesting that male are more likely than their female counterpart to take in high-risk practice, and the p-value, which is less than 0.05, is statistically significant.

## National map for MRI centers- All US states



### Analysis:

The map shows that there is significant variation in MRI center density across the country. We notice that the MRI density is higher in Florida, Providence, New Jersey, and Delaware. Florida has the highest MRI density of 5.7 centers per 100,000 population. MRI scans are used to diagnose a variety of conditions from torn ligaments to tumors. Florida has the highest

percentage (19.0%) of people above the age of 65 years among all US states, which translates to a total of 3.8m people. The elderly population are more likely to need the services provided by MRI scans than the rest of the population. This might explain why there are more MRI centers in Florida.

P.S. During our data preparation, we found that there was 67.8% of data missing (has NaN value). Therefore, High MRI center density in Florida might be due to careless records in the other states.

## Conclusion

From Statistical test for sole proprietor, we notice that the percentage of females who are sole proprietors is higher than the percentage of males who are sole proprietors. Simultaneously, the percentage of females who are partnership are also higher than that of male. The difference in quantity cannot directly reflect the gender difference. There are two reasons. First of all, we only counted the states designated to our group, and our statistics cannot represent the situation of the whole United States. Secondly, from our statistics, we can see that the number of women is much larger than that of men, which also leads to the final percentage difference.

From Statistical test on risk and reward, it is very obvious that females are more inclined to participate in the practices associated with lower risks. Contrarily, males prefer to choose high-income and high-risk occupations. The reasons of this situation are very complex, one of the reasons might be that women have more responsibility to take care of families. Thus, females are trend to choose stable careers. Males prefer high-reward occupations, which correspond to high-risk. This phenomenon is reflected in many industries, but our statistical test just focus on healthcare industry.

From heatmap for MRI centers of US states, we notice that MRI density is relatively smaller in middle US state. On the contrary, MRI density is higher in western and eastern area of US state.

Florida has the highest MRI density of 5.7 centers per 100,000 population, which is closely connected with the large number of elderly populations in Florida. Secondly, the higher MRI density in eastern and western regions is also related to the better economic development of these Regions compared with middle US state.

## Recommendation

The U.S. government needs to pay attention to the different career choices caused by gender differences, in order to prevent the imbalance of the employment market caused by gender ratio.

More MRI centers can be established in the middle regions to improve the local medical and health level. Florida state needs to pay attention to the number of its elderly population to prevent the burden on the health care system caused by too many elderly people.

## References

- [Number of hospitals & patients per day by states](#)
- [US Census Bureau](#)

- Diagnostic Imaging Centers in the US

---

END