Data Science for Social Welfare

Jimmy Hickey

Winona State University

Pi Mu Epsilon

April 13th 2018

WINONA
STATE UNIVERSITY

Data Science: More than a Buzzword

- Gathering
- Managing
- Analyzing
- Communicating

Examples!

Gathering

- Structured vs. Unstructured
- https://data.seattle.gov/Public-Safety/Seattle-Police-Department-911-Incident-Response/3k2p-39jp
- https://usa.ipums.org/usa-action/variables/group
- https://www.thecurrent.org/playlist/2014-01-01/01

Managing

- Cleaning
- Aggregating Data
- Feature Engineering & Dimension Reduction

Cleaning

Data Management

- The data set contained many missing values.
  - Imputed if possible
    - Infer the missing value to be zero by context or extrapolate it from other data
  - Set to "Unknown"
    - Such as in categorical survey data
  - Left missing
    - If filling it in would be misleading or unfounded
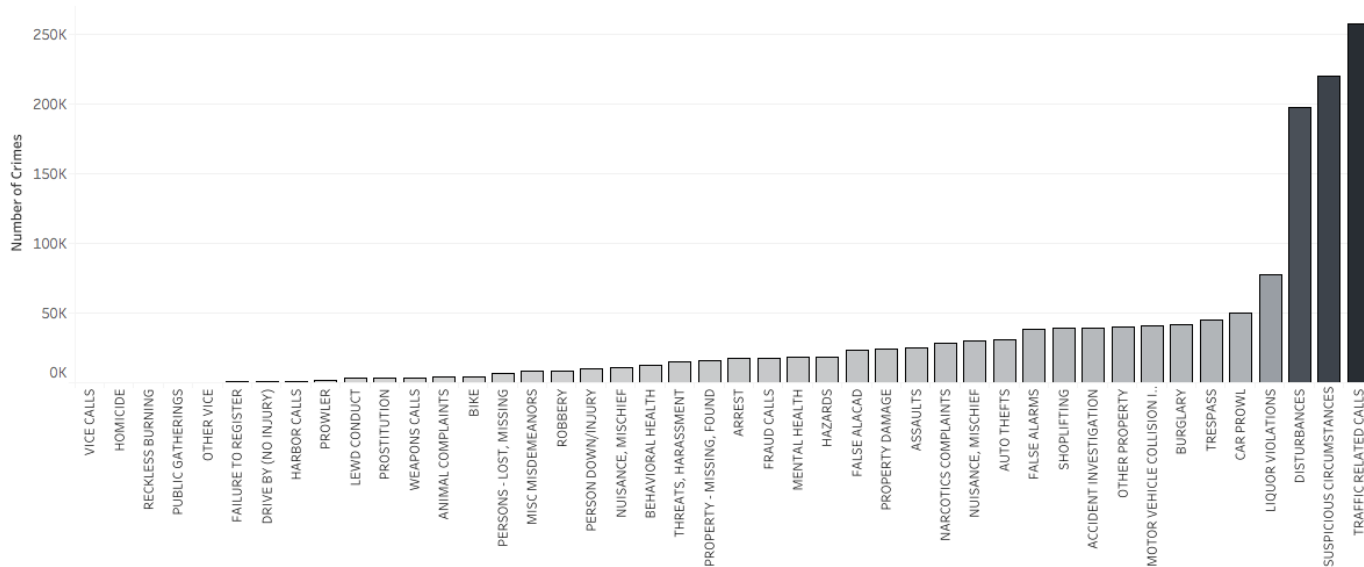- Used models resistant to missing data

Pre–Aggregation

About the Data

We used data from the Seattle Police Department. We had a variety of variables to work with including the location of calls and the type of crime each reported.

A quick glance at types of crimes:

Number of Different Crimes

Post–Aggregation

Our Approach

Our first inclination was to distill the crimes into smaller, more meaningful categories than was provided in the
data set.

These categories are:
- Disturbance/Traffic
- Property
- Miscellaneous
- Health/Drug
- Sexual/Harassment
- Emergency/Violent
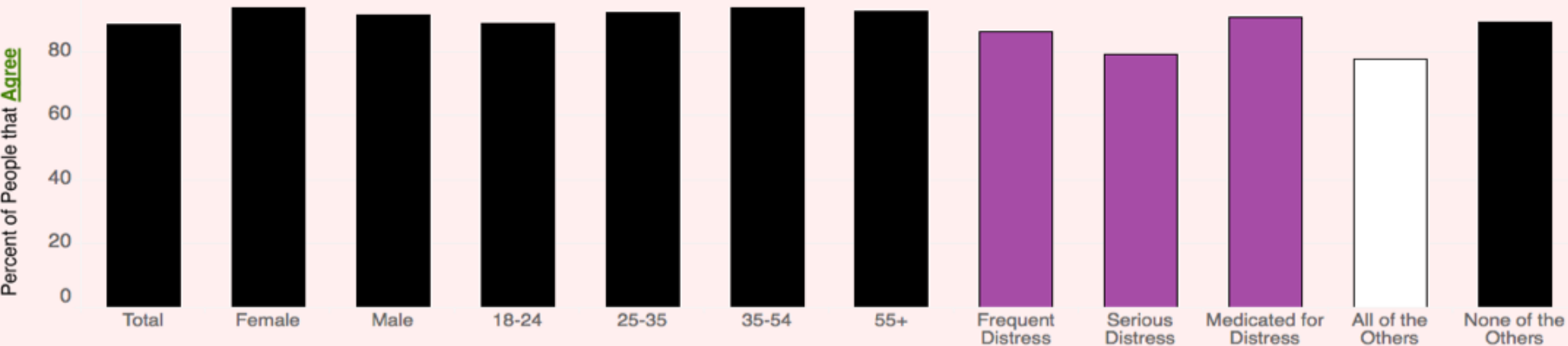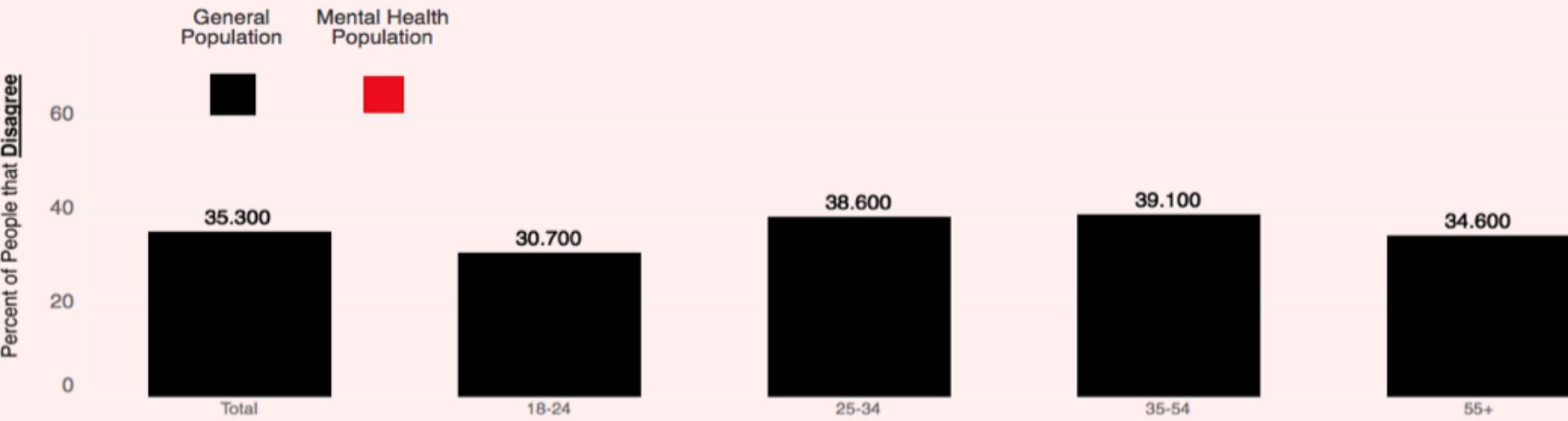
Number of Crimes by Group

Crime Group
- Disturbance/ Traffic
- Property
- Miscellaneous
- Health/Drug
- Sexual/Harassment
- Emergency/ Violent
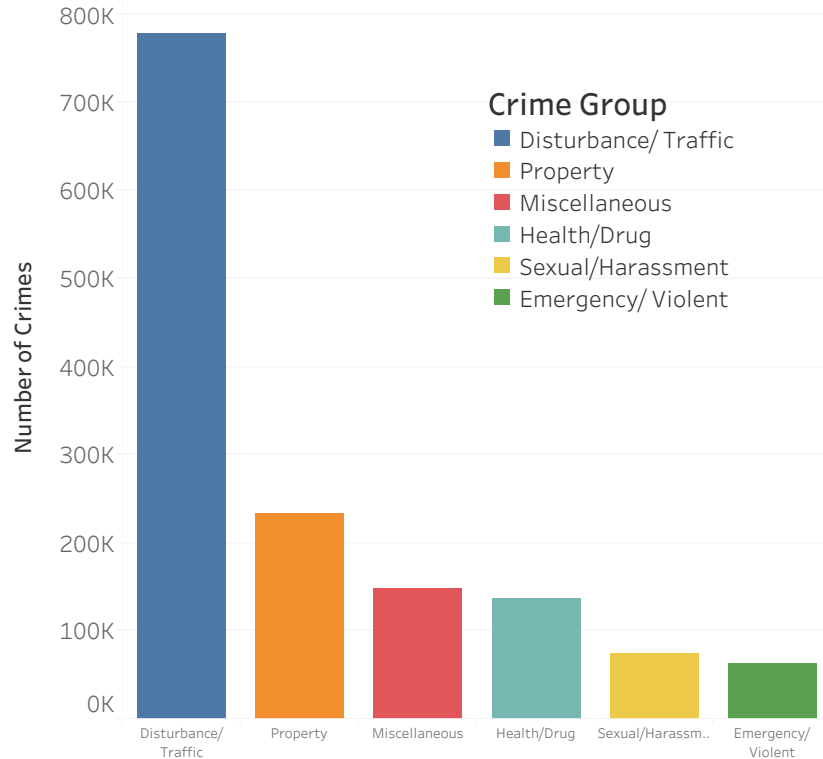
Analyzing

- Visualization
- Modeling
- Both!

Visualizing

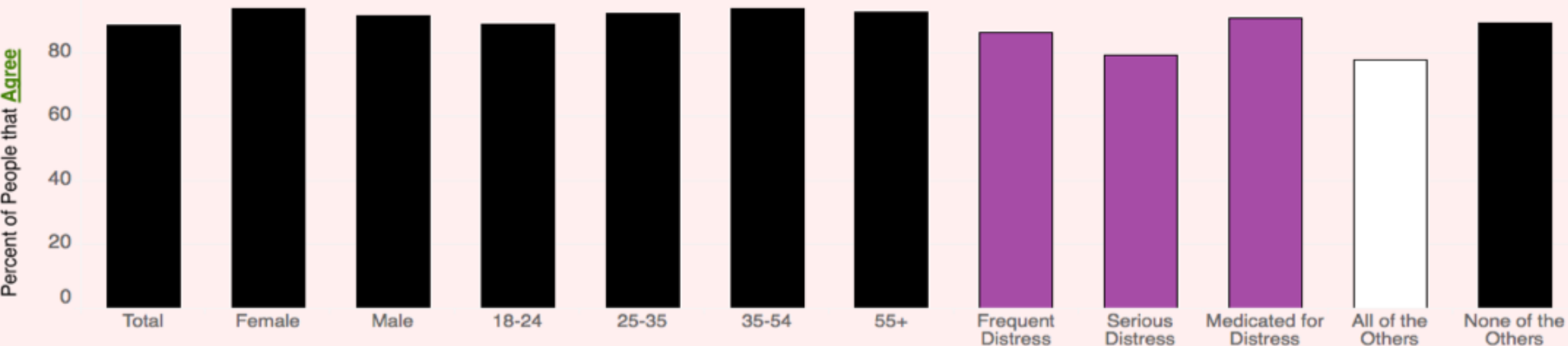# "Treatment can help persons with mental illness lead normal lives."

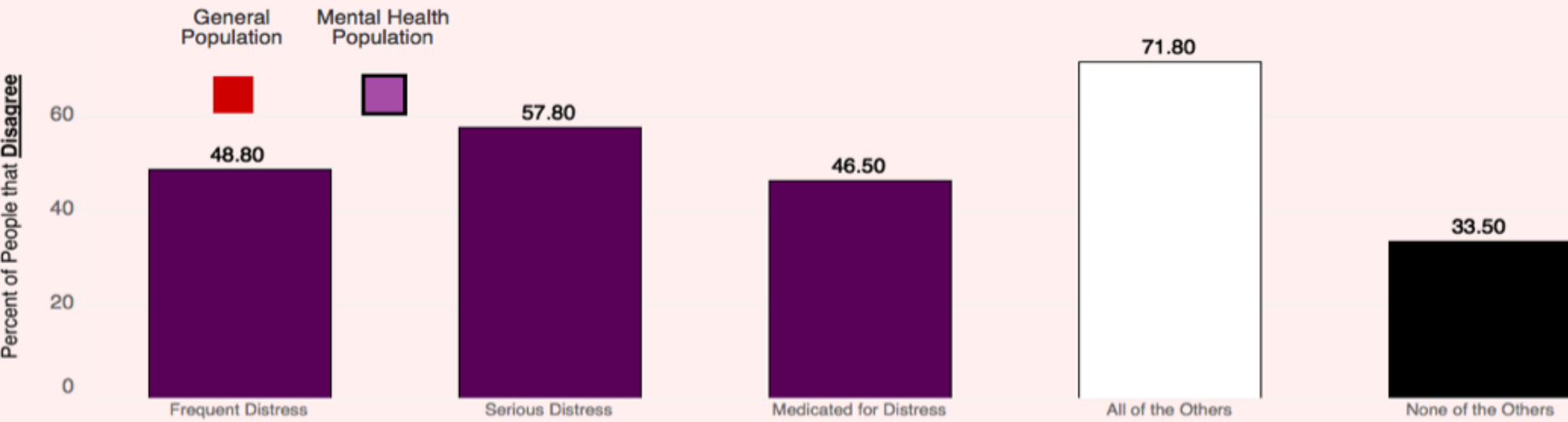Percent of People that **Agree**

| Total | Female | Male | 18-24 | 25-35 | 35-54 | 55+ | Frequent Distress | Serious Distress | Medicated for Distress | All of the Others | None of the Others |

"People are generally caring and sympathetic to persons with mental illness."

General Population — Mental Health Population

Percent of People that **Disagree**

Total	18-24	25-34	35-54	55+
35.300	30.700	38.600	39.100	34.600

"Treatment can help persons with mental illness lead normal lives."

Percent of People that Agree

| Total | Female | Male | 18-24 | 25-35 | 35-54 | 55+ | Frequent Distress | Serious Distress | Medicated for Distress | All of the Others | None of the Others |

"People are generally caring and sympathetic to persons with mental illness."

Percent of People that Disagree

General Population

Mental Health Population

| Frequent Distress | Serious Distress | Medicated for Distress | All of the Others | None of the Others |
| 48.80 | 57.80 | 46.50 | 71.80 | 33.50 |

Emergency Calls

Emergencies During the Week

Emergencies by Sector
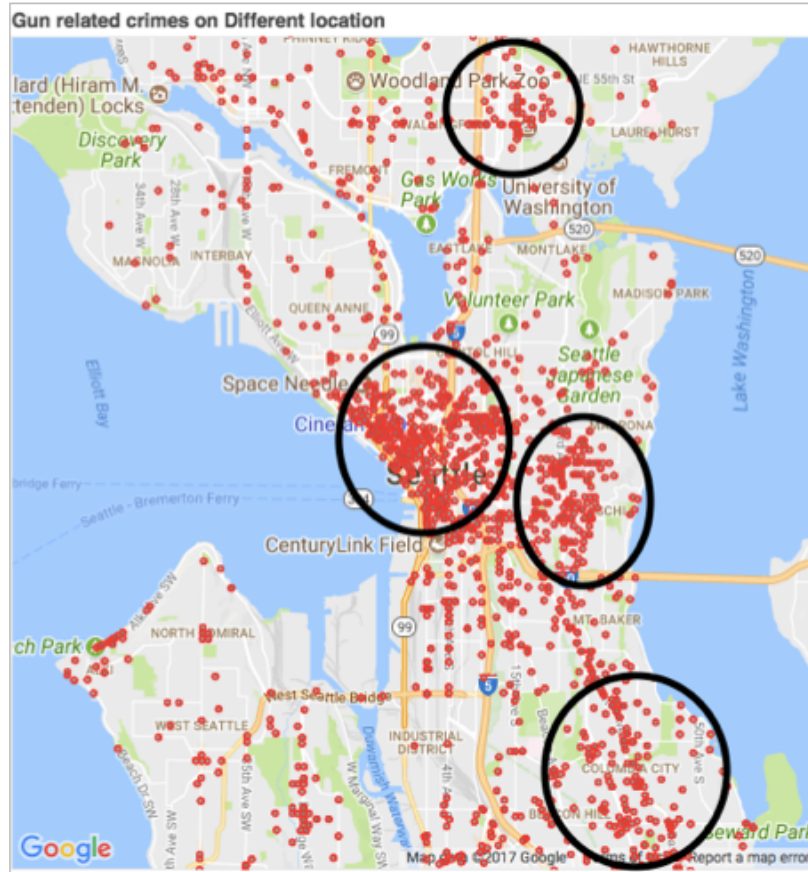
- Similar time patterns hold for emergencies.
- Emergencies are more prevalent in certain sectors.

Geospatial Visualization

Gun related crimes on Different location

Gun related crimes tend to be concentrated in certain areas.

There are four distinct clusters.
- Columbia City
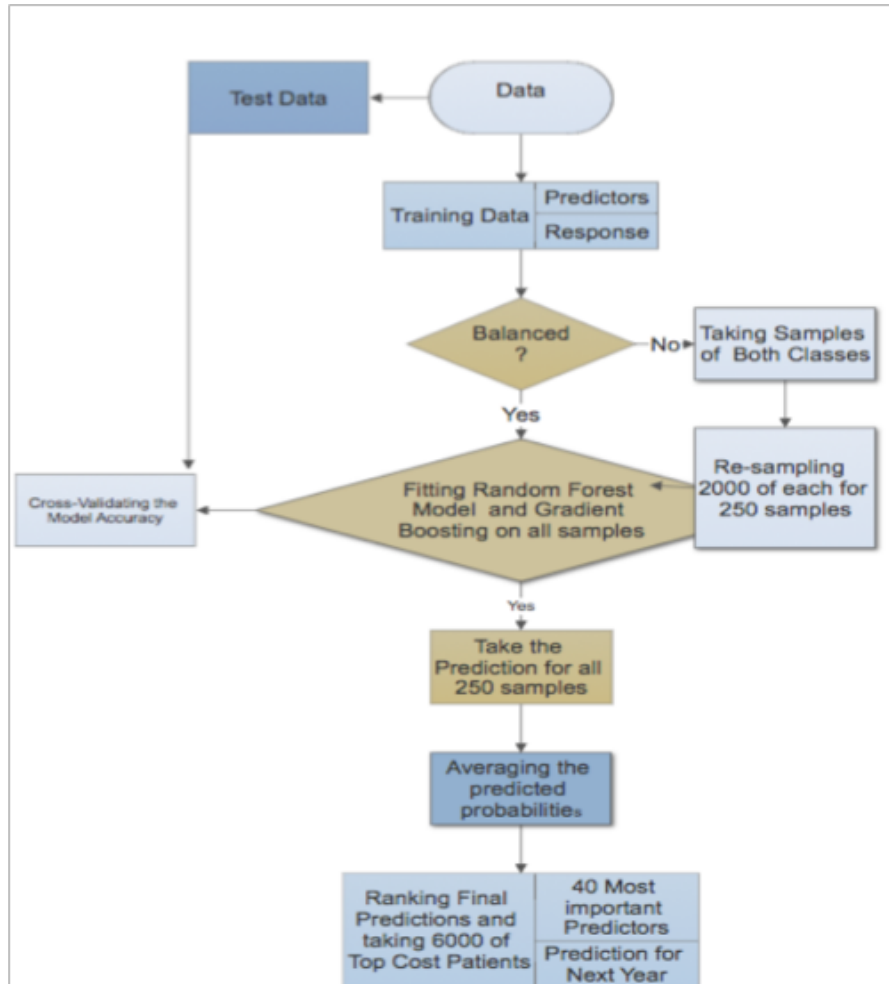- Downtown
- Squire Park
- University District

Modeling

FlowChart of algorithm used for the modeling.

Repeated sampling to balance the data and have decrease its effect.

The Random Forest helps decrease the variance in the data while gradient boosting classifiers to decrease the bias.

Model properly predicts around 55% of 6000 patients.

19

Both!

Variables that are Most Explanatory of Top 6000

Communicating

- Visualizations

- Understanding your audience

- Recommendations

Using Data to Better Understand Mental Health

Jimmy Hickey

Using Medical Data to Predict Future Patient Expenditures

Will Diedrick, Jimmy Hickey, Akif Khan,
Kapil Khanal, Sean Wittenberg

Brad Erickson, Uzma Ghazanfar, Brad Hennes, Jimmy Hickey

TEAM 5

Midwest Undergraduate Data
Analytics Competition

Winona State University | www.MUDAC.org

Conclusion

- The predictions from our model will help the employer understand the characteristics of diabetic employees that are likely to be high cost in the future.
- This insight will help allow the employer to identify these employees and focus on improving the health of these patients.
- Successfully identifying these employees will both reduce costs for the employer but also benefits the potentially high cost employees.

26

Improving one's mental health

These problems in society's treatment of those with mental health problems are not going to change overnight.

There are still many ways that one who feels stigmatized can help themself.

Learn
Online Therapy Sleep Change Environment
Animal Therapy Time Outdoors
Exercise Managing Vices Hobbies
Organize Socialize Meditation Talk
Relax Volunteering Diet Therapy
Support Network

People with **mental health problems** sleep less than **those without**.
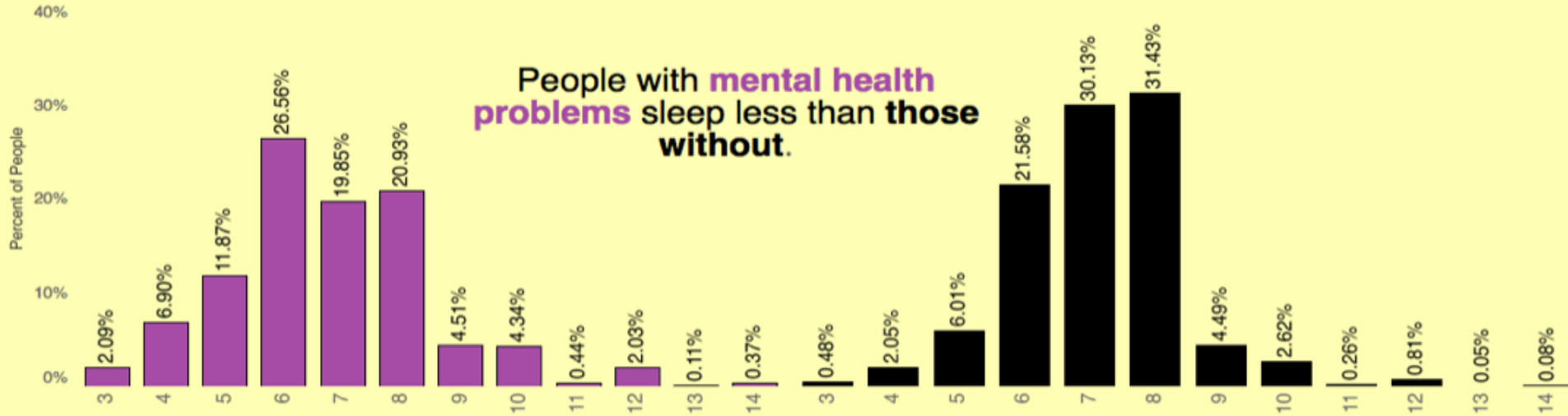
Insight
Emergencies are highest in sector K around 7 P.M. during the late summer.

Gun related crimes tend to be clustered in certain areas.

Recommendations
The public should be made aware of the increased danger during these times and in these locations. A PSA campaign could disseminate this information effectively.

The Seattle Police Department should increase patrols in these hotspots during the most active times.

Acknowledgements

- Will Diedrick
- Brad Erickson
- Uzma Ghanzanfar
- Brad Hennes
- Akif Khan
- Kapil Khanal
- Sean Wittenberg

Questions?