

# CSE 180

## Introduction to Data Science

Instructor:

Ryan Maas

Skill it Up creator and consultant:

Bob Boiko

Teaching Assistants:

Ryan Fok  
Samuel Gehman  
Kevin Kang  
Trevor Perrier  
Andrew Wei

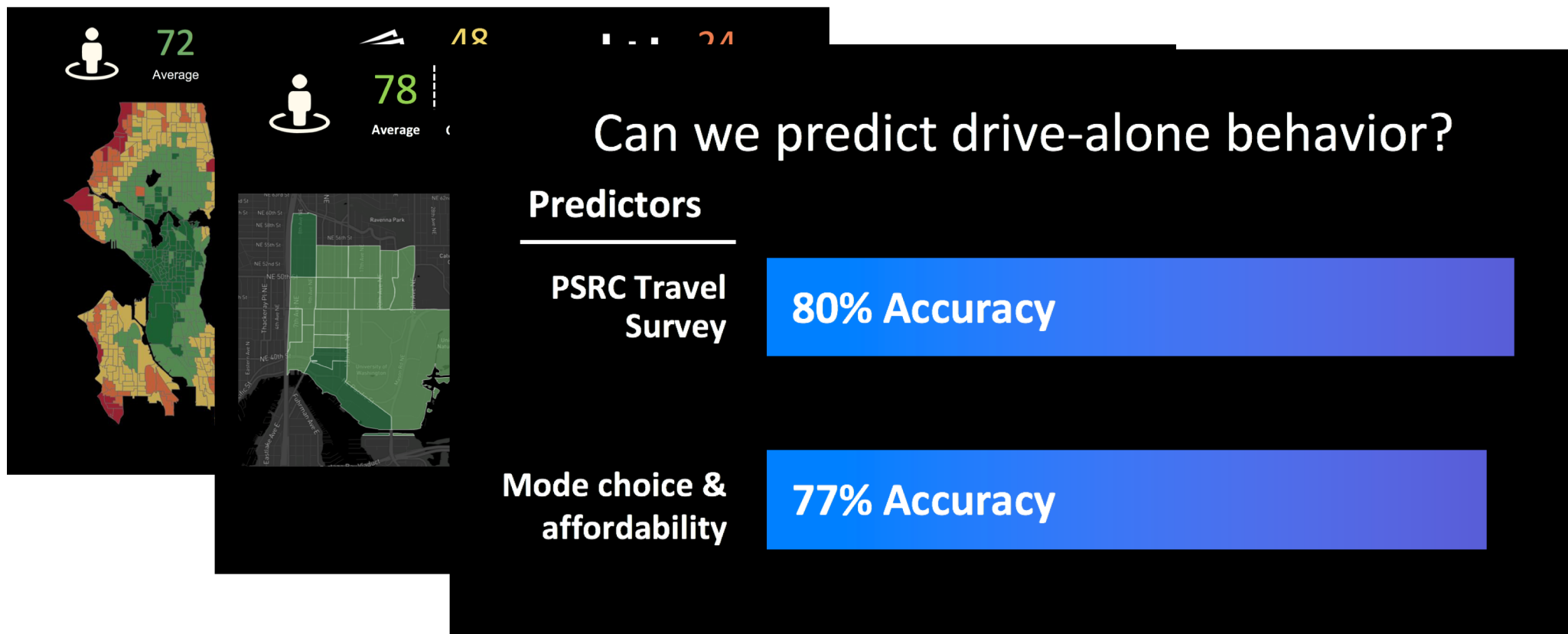


# Outline

1. What is data science?
2. Python for data science
3. Statistics for data science

# What is data science?

Drawing useful conclusions from large and diverse data sets through **exploration, inference, and prediction.**



# What Python skills do we need?

- Data exploration/processing:
  - The Python scientific stack
  - numpy, scipy, pandas, matplotlib
- Formatting and presenting results:
  - JuPyter notebooks



# What statistics skills do we need?

- Basics:
  - Estimation
  - Hypothesis testing
  - Inference
- Applied statistics:
  - Classification



# What other skills do we need?

- Combining statistics and programming:
  - Machine learning
- Communicating your results
  - Jupyter Lab and notebooks
- Adapting to different types of data
  - Evaluating data sets
  - Cleaning messy data