

# Foundations of Data Science with R

Jameson Watts, Ph.D.

# Agenda

1. Course Overview and Expectations
2. Example Analysis of Wine Prices

## Course Overview and Expectations

## Contact Information

- ▶ Email: [jwatts@willamette.edu](mailto:jwatts@willamette.edu)
- ▶ Office Hours: after class
- ▶ Appointments: [jamesonwatts.youcanbook.me](https://jamesonwatts.youcanbook.me) (Skype or phone call)

# Class Materials

- ▶ Base R
- ▶ RStudio 1.2
- ▶ R for Data Science
- ▶ DataCamp Classroom

Other resources:

- ▶ <https://twitter.com/r4dscommunity>
- ▶ <https://bookdown.org/yihui/rmarkdown>
- ▶ <http://google.com>

# Reading the Course Outline

- ▶ Class Topics
  - ▶ Subjects I plan to cover during that day's lecture
- ▶ Reading and Assignments
  - ▶ DCC: assignments in the DataCamp Classroom
  - ▶ R4ds: chapters to read in the online textbook

# Assignments

- ▶ DataCamp homework assignments (25%)
- ▶ Midterm exams (50%)
- ▶ Final Presentations and Report (25%)

# Course Policies and Expectations

- ▶ Name tents
- ▶ Collaboration
- ▶ Late work
- ▶ Effort
  - ▶ 2-4 hours outside of class each week
  - ▶ struggle, Google, struggle, doh!
  - ▶ start with the basics. . . ramp up very fast



## Analysis of wine prices

# Dataset

```
library(tidyverse)
```

```
## -- Attaching packages -----
```

```
## v ggplot2 3.2.1      v purrr   0.3.2  
## v tibble  2.1.3      v dplyr   0.8.3  
## v tidyr   0.8.3      v stringr 1.4.0  
## v readr   1.3.1      v forcats 0.4.0
```

```
## -- Conflicts -----
```

```
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag()     masks stats::lag()
```

```
summary(cars)
```

```
##           speed           dist  
##  Min.      : 4.0    Min.      : 2.00  
## 1st Qu.:12.0    1st Qu.: 26.00  
## Median :15.0    Median : 36.00  
## Mean   :15.4    Mean    : 42.98
```

## Slide with Plot

