

Project Proposal - Tennis Trio

Due Friday, October 9, 11:59 PM

Tennis Trio: Hamilton Murrah, Kellyn McDonald, Naima Turbes

```
library(tidyverse)
wine_data <- read_csv("data/winemag-data_first150k.csv")
wine_data

## # A tibble: 150,930 x 11
##       X1 country description designation points price province region_1 region_2
##   <dbl> <chr>   <chr>         <chr>         <dbl> <dbl> <chr>   <chr>   <chr>
## 1     0 US      This treme~ Martha's V~    96   235 Califor~ Napa Va~ Napa
## 2     1 Spain   Ripe aroma~ Carodorum ~    96   110 Norther~ Toro    <NA>
## 3     2 US      Mac Watson~ Special Se~    96    90 Califor~ Knights~ Sonoma
## 4     3 US      This spent~ Reserve      96    65 Oregon  Willame~ Willame~
## 5     4 France   This is th~ La Brûlade    95    66 Provence Bandol  <NA>
## 6     5 Spain   Deep, dens~ Numanthia     95    73 Norther~ Toro    <NA>
## 7     6 Spain   Slightly g~ San Román     95    65 Norther~ Toro    <NA>
## 8     7 Spain   Lush cedar~ Carodorum ~    95   110 Norther~ Toro    <NA>
## 9     8 US      This re-na~ Silice      95    65 Oregon  Chehale~ Willame~
## 10    9 US      The produc~ Gap's Crow~    95    60 Califor~ Sonoma ~ Sonoma
## # ... with 150,920 more rows, and 2 more variables: variety <chr>, winery <chr>
```

Introduction

Our primary research question is as follows: What is the relationship between the best wines and where these wines from? Our hypothesis is that the best wines are from France and Italy because we know that this is where most wines come from. We also hypothesize that these higher rated wines will be more expensive.

Data Description

This data was collected by scraping data from wine enthusiasts.

Glimpse of data