

Proposal

STA 210 - Project

Ginger and Stats - Aimy Wen, Rakshita Ramakrisna, Nathan Nguyen, Bryan Pan

```
library(tidyverse)
library(tidymodels)

chocolate <- read_csv("../data/chocolate.csv")

glimpse(chocolate)
```

Rows: 2,530

Columns: 11

\$...1	<dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12~
\$ ref	<dbl> 2454, 2458, 2454, 2542, 2546, 2546, 2~
\$ company_manufacturer	<chr> "5150", "5150", "5150", "5150", "5150~
\$ company_location	<chr> "U.S.A.", "U.S.A.", "U.S.A.", "U.S.A.~
\$ review_date	<dbl> 2019, 2019, 2019, 2021, 2021, 2021, 2~
\$ country_of_bean_origin	<chr> "Tanzania", "Dominican Republic", "Ma~
\$ specific_bean_origin_or_bar_name	<chr> "Kokoa Kamili, batch 1", "Zorzal, bat~
\$ cocoa_percent	<chr> "76%", "76%", "76%", "68%", "72%", "8~
\$ ingredients	<chr> "3- B,S,C", "3- B,S,C", "3- B,S,C", "~
\$ most_memorable_characteristics	<chr> "rich cocoa, fatty, bready", "cocoa, ~
\$ rating	<dbl> 3.25, 3.50, 3.75, 3.00, 3.00, 3.25, 3~

Introduction

...

Data description

...

Analysis approach

...

Data dictionary

The data dictionary can be found [here](#).