Data Science Project with App Store Dataset

Harshdeep Kahlon

12/4/2019

1. Abstract

The dataset found here contains more than 7,000 details about apps that are currently listed on the Apple iOS App Store. By investigating this data, we can potentially find interesting observations that could assist companies and developers to further grow their mobile applications.

2. Problem Descriptions and Objectives

The goal of this project is to analyze the App Store data and discover if genre, price, and/or content rating are vital in determining the success of an app. Once the attributes have been found, this project will try to classify whether or not a successful app needs a high user rating. This project will also explore the relationship between an app's size in megabytes and the genre of it. Finally, this project will compare free apps to paid apps ($0.99 and above) and find any major differences.

3. Data Description

- X - quantitative
- id - quantitative
- track_name - categorical
- size_bytes - quantitative
- currency - categorical
- price - quantitative
- rating_count_tot - quantitative
- rating_count_ver - quantitative
- user_rating - quantitative
- user_rating_ver - quantitative
- ver - categorical
- cont_rating - categorical
- prime_genre - categorical
- sup_devices.num - quantitative
- ipadSc_urls.num - quantitative
- lang.num - quantitative
- vpp_lic - quantitative

4. Loading and Formatting Data

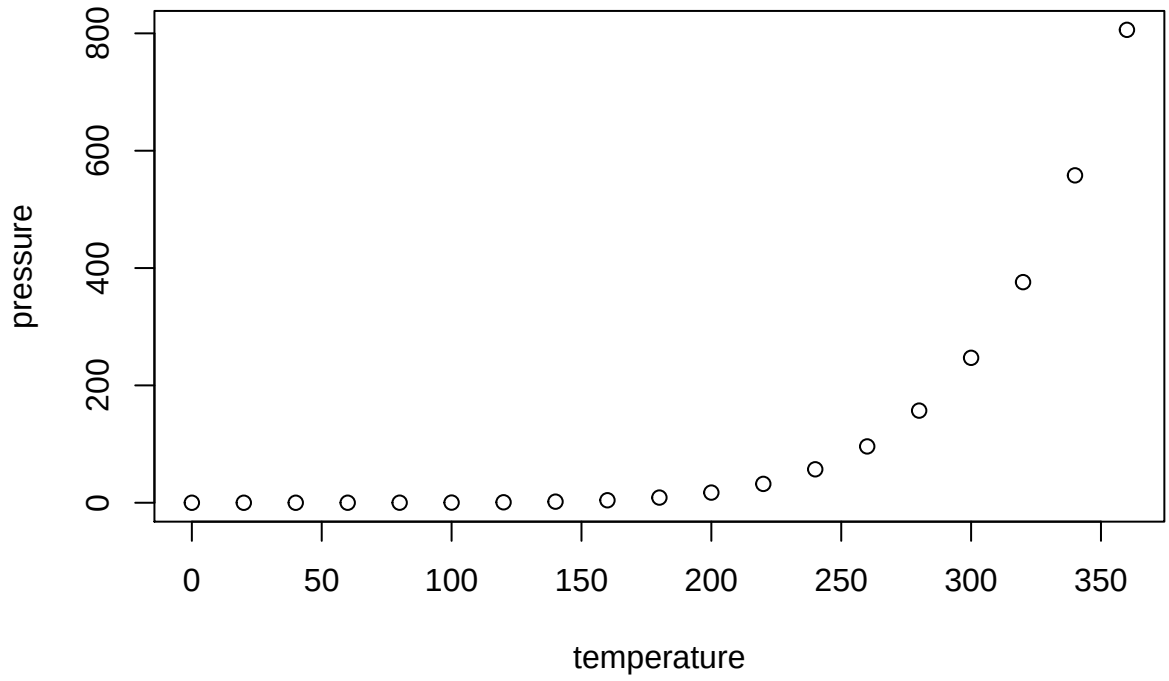
```
summary(cars)
```

```
##      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
```

```
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.