Week 12 Project Submission

Heng Javier

2023-10-25

Week 9 Questions:

(1) What is the topic that you have finalized? (Answer in 1 or 2 sentences).

I will be doing an analysis of Taylor Swift's music/albums on Spotify.Specifically, I will be focusing on aspects or her music such as tempo, valence and time signature.

(2) What are the data sources that you have curated so far? (Answer 1 or 2 sentences).

I will be using the Taylor Swift dataset from Tidy Tuesday. Specifically the taylor_all_songs dataset.

Week 10 Questions:

(1) What is the question that you are going to answer? (Answer: One sentence that ends with a question mark that could act like the title of your data story),

What is Taylor Swift's music like? And how has her music changed over the years?

(2) Why is this an important question? (Answer: 3 sentences, each of which has some evidence, e.g., "According to the United Nations. . . " to justify why the question you have chosen is important),

In August 2023, Taylor Swift was the first female artist to reach 100 million monthly listeners. Taylor Swift has a $740 million net worth as of June 2023 (according to Forbes). Analysing the stats of her music will help other musicians to figure out the ingredients in her music to cook up hits just like her, and figuring out how her music changed can give us some insights to how she has changed as a celebrity.

(3) Which rows and columns of the dataset will be used to answer this question? (Answer: Actual names of the variables in the dataset that you plan to use).

Columns I will use include track_name, danceability, energy, key, loudness, mode, speechiness,acousticness, instrumentalness, liveness, tempo, explicit, key_mode, track_release. All variables are defined as per https://github.com/rfordatascience/tidytuesday/blob/master/data/2023/2023-10-17/readme.md.

As for rows used, I will use every row! However, data will be cleaned in the event of NA data. #(1) List the visualizations that you are going to use in your project (Answer: What are the variables that you are going to plot? How will it answer your larger question?) I will plot the following graphs: -Graph of danceability, energy, tempo against track_release. This will allow us to see if Swift's music has become more energetic or slower and more melancholic over the years.

-Graph of danceability against energy, tempo, loudness. Does danceability mean more energy and loud music? It is important to know the difference between these variables to properly understand the data.

-Graph of acousticness/liveness against track_release Did Swift have more acoustic or live tracks at any point in time? This graph will help us to see more.

-key_mode against track_name To find out if Swift has a favourite key for her music.

#(2) How do you plan to make it interactive? (Answer: features of ggplot2/shiny/markdown do you plan to use to make the story interactive) Using Shiny, I can include some ggplots which can be changed around when clicking a different tab in a dropdown menu. Slider Youtube videos of Taylor's highest scoring songs Radio button to display a song that scores high on a particular domain if the user likes that genre of music.

#(3) What concepts incorporated in your project were taught in the course and which ones were self-learnt? (Answer: Create a table with topics in one column and Weeks in the other to indicate which concept taught in which week is being used. Leave the entry of the Week column empty for self-learnt concepts)

```
# create matrix with 4 columns and 4 rows
data= matrix(c("ggplot","2&7","Tidy data, selecting, filtering and arranging", 4, "Functions", 5, "Shiny
# specify the column names and row names of matrix
colnames(data) = c('concept','week')
rownames(data) <- c(1:5)

# assign to table
final=as.table(data)

# display
final
```

```
##   concept                                          week
## 1 ggplot                                           2&7
## 2 Tidy data, selecting, filtering and arranging 4
## 3 Functions                                        5
## 4 Shiny                                            8
## 5 if else                                          0
```

Week 12 Questions:

#(1)Include the challenges and errors that you faced and how you overcame them (if any) A lot of code was not taught and interactive features required me to think out of the box. I first looked for code online, then asked my TAs to help me make sense of the code before applying it.

```r
# create matrix with 4 columns and 4 rows
data= matrix(c(1:16), ncol=4, byrow=TRUE)

# specify the column names and row names of matrix
colnames(data) = c('col1','col2','col3','col4')
rownames(data) <- c('row1','row2','row3','row4')

# assign to table
final=as.table(data)

# display
final
```

```
##      col1 col2 col3 col4
## row1    1    2    3    4
## row2    5    6    7    8
## row3    9   10   11   12
## row4   13   14   15   16
```

Dataset used: taylor_all_songs

```r
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.3     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.3     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
install.packages("tidytuesdayR",repos = "http://cran.us.r-project.org")
```

```
## package 'tidytuesdayR' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\hengj\AppData\Local\Temp\RtmpUxGIHj\downloaded_packages
```

```r
library(tidytuesdayR)

tuesdata <- tidytuesdayR::tt_load('2023-10-17')
```

```
## --- Compiling #TidyTuesday Information for 2023-10-17 ----
## --- There are 3 files available ---
## --- Starting Download ---
```

```
##
##  Downloading file 1 of 3: `taylor_album_songs.csv`
##  Downloading file 2 of 3: `taylor_all_songs.csv`
##  Downloading file 3 of 3: `taylor_albums.csv`

## --- Download complete ---
```

```
taylor_all_songs <- tuesdata$taylor_all_songs


taylor_all_songs
```

```
## # A tibble: 274 x 29
##    album_name   ep    album_release track_number track_name     artist featuring
##    <chr>        <lgl> <date>               <dbl> <chr>          <chr>  <chr>
##  1 Taylor Swift FALSE 2006-10-24               1 Tim McGraw     Taylo~ <NA>
##  2 Taylor Swift FALSE 2006-10-24               2 Picture To Bu~ Taylo~ <NA>
##  3 Taylor Swift FALSE 2006-10-24               3 Teardrops On ~ Taylo~ <NA>
##  4 Taylor Swift FALSE 2006-10-24               4 A Place In Th~ Taylo~ <NA>
##  5 Taylor Swift FALSE 2006-10-24               5 Cold As You    Taylo~ <NA>
##  6 Taylor Swift FALSE 2006-10-24               6 The Outside    Taylo~ <NA>
##  7 Taylor Swift FALSE 2006-10-24               7 Tied Together~ Taylo~ <NA>
##  8 Taylor Swift FALSE 2006-10-24               8 Stay Beautiful Taylo~ <NA>
##  9 Taylor Swift FALSE 2006-10-24               9 Should've Sai~ Taylo~ <NA>
## 10 Taylor Swift FALSE 2006-10-24              10 Mary's Song (~ Taylo~ <NA>
## # i 264 more rows
## # i 22 more variables: bonus_track <lgl>, promotional_release <date>,
## #   single_release <date>, track_release <date>, danceability <dbl>,
## #   energy <dbl>, key <dbl>, loudness <dbl>, mode <dbl>, speechiness <dbl>,
## #   acousticness <dbl>, instrumentalness <dbl>, liveness <dbl>, valence <dbl>,
## #   tempo <dbl>, time_signature <dbl>, duration_ms <dbl>, explicit <lgl>,
## #   key_name <chr>, mode_name <chr>, key_mode <chr>, lyrics <lgl>
```