

Homework I

Cagri Gokpunar

10/29/2021

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.3.5      v purrr 0.3.4
## v tibble 3.1.5       v dplyr 1.0.7
## v tidyr 1.1.4        v stringr 1.4.0
## v readr 2.0.2        v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()

library(magrittr)

##
## Attaching package: 'magrittr'

## The following object is masked from 'package:purrr':
##
## set_names

## The following object is masked from 'package:tidyr':
##
## extract

maindf <- read.csv("Revenue_Chart_Full_Data_data.csv")
df <- read.csv("Revenue_Chart_Full_Data_data.csv")

df %<>% select("Year..copy.", "Format", "Value..Actual.")

df %<>% rename(Year= "Year..copy.", Value = "Value..Actual.")

dfYear <- filter(df, Year > 2017 )

dfYear %<>% arrange(Year)

dfYear2018 <- filter(dfYear, Year==2018)

sumdfYear2018 <- sum(dfYear2018$Value)
dfYear2018

##      Year                Format      Value
## 1  2018                Vinyl Single 5.720601e+00
## 2  2018      SoundExchange Distributions 9.528000e+02
## 3  2018      Ringtones & Ringbacks 2.495786e+01
## 4  2018                      SACD 8.596997e-01
## 5  2018      Download Music Video 2.220811e+00
```

```
## 6 2018          Paid Subscription 4.614019e+03
## 7 2018          Kiosk 1.965046e+00
## 8 2018          DVD Audio 3.288671e-01
## 9 2018          CD 6.958136e+02
## 10 2018         Music Video (Physical) 2.841531e+01
## 11 2018         Other Ad-Supported Streaming 2.081769e+02
## 12 2018         CD Single 1.580415e-02
## 13 2018         Other Digital 1.983957e+01
## 14 2018         Limited Tier Paid Subscription 7.405133e+02
## 15 2018         Download Single 4.898662e+02
## 16 2018         LP/EP 4.192239e+02
## 17 2018         Download Album 4.952583e+02
## 18 2018         Synchronization 2.855135e+02
## 19 2018 On-Demand Streaming (Ad-Supported) 7.526890e+02
```

```
dfYear2019 <- filter(dfYear, Year==2019)
```

```
sumdfYear2019 <- sum(dfYear2019$Value)
dfYear2019
```

##	Year	Format	Value
## 1	2019	Synchronization	2.811127e+02
## 2	2019	CD Single	6.255867e-02
## 3	2019	CD	6.306914e+02
## 4	2019	Paid Subscription	6.115172e+03
## 5	2019	Kiosk	1.586471e+00
## 6	2019	DVD Audio	1.279680e+00
## 7	2019	SACD	4.333023e-01
## 8	2019	Other Digital	2.151245e+01
## 9	2019	Other Ad-Supported Streaming	2.072994e+02
## 10	2019	Download Music Video	1.855022e+00
## 11	2019	Download Album	3.687779e+02
## 12	2019	Vinyl Single	6.728344e+00
## 13	2019	SoundExchange Distributions	9.082000e+02
## 14	2019	On-Demand Streaming (Ad-Supported)	1.013055e+03
## 15	2019	Limited Tier Paid Subscription	6.382262e+02
## 16	2019	Music Video (Physical)	2.583715e+01
## 17	2019	LP/EP	4.795284e+02
## 18	2019	Download Single	4.083530e+02
## 19	2019	Ringtones & Ringbacks	2.064132e+01

```
dfYear2020 <- filter(dfYear, Year==2020)
```

```
sumdfYear2020 <- sum(dfYear2020$Value)
dfYear2020
```

##	Year	Format	Value
## 1	2020	Download Single	312.8190580
## 2	2020	Vinyl Single	6.3172200
## 3	2020	Ringtones & Ringbacks	20.2396063
## 4	2020	Other Ad-Supported Streaming	211.2476148
## 5	2020	CD	483.3405330
## 6	2020	Kiosk	1.1768843
## 7	2020	Paid Subscription	7009.1654559
## 8	2020	Limited Tier Paid Subscription	723.6146792
## 9	2020	SoundExchange Distributions	947.4000000

```
## 10 2020 Download Album 319.5010023
## 11 2020 Synchronization 265.2360669
## 12 2020 LP/EP 619.6100603
## 13 2020 CD Single 0.3703540
## 14 2020 Other Digital 18.8902642
## 15 2020 DVD Audio 1.8437452
## 16 2020 Music Video (Physical) 27.4266245
## 17 2020 SACD 0.2366725
## 18 2020 On-Demand Streaming (Ad-Supported) 1183.1213597
## 19 2020 Download Music Video 1.7942255
```

```
sumdfYear2018
```

```
## [1] 9738.197
```

```
sumdfYear2019
```

```
## [1] 11130.35
```

```
sumdfYear2020
```

```
## [1] 12153.35
```

```
streaming <- c("Paid Subscription",
"On-Demand Streaming (Ad-Supported)", "Other Ad-Supported Streaming", "SoundExchange Distributions",
"Limited Tier Paid Subscription")
dfYear2020streaming <- filter(dfYear2020, Format %in% streaming)
dfYear2020streaming
```

```
## Year Format Value
## 1 2020 Other Ad-Supported Streaming 211.2476
## 2 2020 Paid Subscription 7009.1655
## 3 2020 Limited Tier Paid Subscription 723.6147
## 4 2020 SoundExchange Distributions 947.4000
## 5 2020 On-Demand Streaming (Ad-Supported) 1183.1214
```

```
sumdfYear2020streaming <- sum(dfYear2020streaming$Value)
sumdfYear2020streaming
```

```
## [1] 10074.55
```

```
synchronization <- c("Synchronization")
physical <- c("CD", "LP/EP")
digitalDownloads <- c("Download Album", "Download Single")
```

```
dfYear2020synchronization <- filter(dfYear2020, Format %in% synchronization)
dfYear2020physical <- filter(dfYear2020, Format %in% physical)
dfYear2020digitalDownloads <- filter(dfYear2020, Format %in% digitalDownloads)
```

```
sumdfYear2020synchronization <- sum(dfYear2020synchronization$Value)
sumdfYear2020physical <- sum(dfYear2020physical$Value)
sumdfYear2020digitalDownloads <- sum(dfYear2020digitalDownloads$Value)
sumdfYear2020synchronization
```

```
## [1] 265.2361
```

```
sumdfYear2020physical
```

```
## [1] 1102.951
```

```
sumdfYear2020digitalDownloads
```

```
## [1] 632.3201
```

```
sumdfYear2020other=sumdfYear2020-(sumdfYear2020streaming+sumdfYear2020synchronization+sumdfYear2020physical+sumdfYear2020digitalDownloads)
sumdfYear2020other
```

```
## [1] 78.2956
```

```
round((sumdfYear2020streaming/sumdfYear2020)*100, 2)
```

```
## [1] 82.9
```

```
round((sumdfYear2020synchronization/sumdfYear2020)*100, 2)
```

```
## [1] 2.18
```

```
round((sumdfYear2020physical/sumdfYear2020)*100, 2)
```

```
## [1] 9.08
```

```
round((sumdfYear2020digitalDownloads/sumdfYear2020)*100, 2)
```

```
## [1] 5.2
```

```
SoundExchangeDistributions <-c("SoundExchange Distributions")
```

```
OtherAdSupportedStreaming <-c("Other Ad-Supported Streaming")
```

```
dfSoundExchangeDistributions <- filter(dfYear, Format %in% SoundExchangeDistributions)
```

```
dfOtherAdSupportedStreaming <- filter(dfYear, Format %in% OtherAdSupportedStreaming)
```

```
graphData<-rbind(dfSoundExchangeDistributions,dfOtherAdSupportedStreaming)
```

```
ggplot(graphData, aes(fill=Format, y=Value, x=Year)) +  
  geom_bar(position="stack", stat="identity")
```

