SteamSpy

Jie Hu

June 2, 2016

SteamSpy API Data

The data I get comes from SteamSpy API with request name "top100in2weeks" Click here (http://steamspy.com/about) to enter SteamSpy API website. (As mentioned by author, sometimes Valve Web API servers do not work as intendedSergey, this API may have missing data or wrong information)

The API accept GET method and outputs JSON data format.

```
top100data <- read.csv("test.csv", sep=',')
# head(top100data)
```

Here I just make initial comparison among these games on Total Players and Average Play Time in last 2 weeks.

```
# Plot Top 10 games with most players from beginning

newdata_rank <- top100data[order(-top100data$TotalPlayer),]

newdata_rank$Name <- factor(newdata_rank$Name, levels = newdata_rank[order(newd
ata_rank$TotalPlayer),"Name"])

library(ggplot2)
library(gridExtra)
ggplot(data=newdata_rank[1:10,], aes(x=factor(Name), y=TotalPlayer)) +
  geom_bar(colour="black", stat="identity",fill="#428fa6") +
  guides(fill=FALSE) +
  geom_text(aes(label=paste(as.character(round(TotalPlayer/1000000, digits =
2)),"M",sep="")), position=position_dodge(width=1), hjust = -0.1,vjust=0.5, siz
e = 5) +
  theme(plot.title = element_text(size=20),
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_blank(),
        panel.background = element_blank(),
        axis.text.y = element_text(face="bold", color="#993333",
                        size=14)) +
  scale_y_continuous(limits = c(0, 100000000), breaks = NULL) +
  labs(title = "Total Player") +
  xlab("Game Names")+ylab("") +
  coord_flip()
```

Total Player

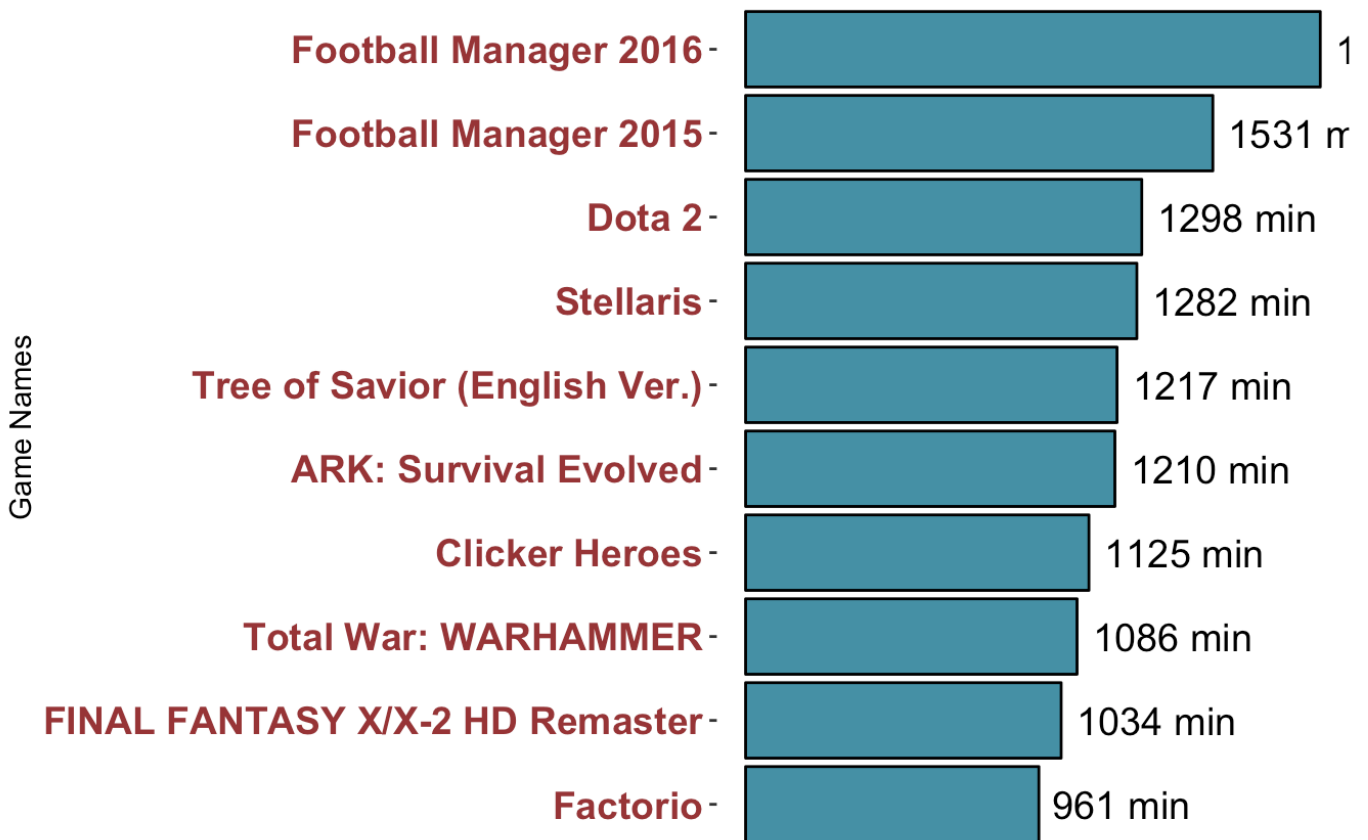| Game	Players
Dota 2	72.17M
Team Fortress 2	33.57M
Counter-Strike: Global Offensive	21.2M
Unturned	17.08M
Left 4 Dead 2	12.27M
Counter-Strike: Source	10.98M
Garry's Mod	10.57M
The Elder Scrolls V: Skyrim	9.97M
Counter-Strike	8.75M
Sid Meier's Civilization V	8.26M

Game Names

```
# Plot Top 10 games with players spending most time last 2 weeks

newdata_avg <- top100data[order(-top100data$AvgPlayTimeIn2weeks),]

newdata_avg$Name <- factor(newdata_avg$Name, levels = newdata_avg[order(newdata
_avg$AvgPlayTimeIn2weeks),"Name"])

ggplot(data=newdata_avg[1:10,], aes(x=factor(Name), y=AvgPlayTimeIn2weeks)) +
  geom_bar(colour="black", stat="identity",fill="#428fa6") +
  guides(fill=FALSE) +
  geom_text(aes(label=paste(as.character(AvgPlayTimeIn2weeks)," min",sep="")),
position=position_dodge(width=1), hjust = -0.1,vjust=0.5, size = 5) +
  theme(plot.title = element_text(size=20),
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank(),
        panel.border = element_blank(),
        panel.background = element_blank(),
        axis.text.y = element_text(face="bold", color="#993333",
                        size=14)) +
  scale_y_continuous(breaks=NULL) +
  labs(title = "Average Play Time in last 2 weeks") +
  xlab("Game Names")+ylab("") + coord_flip()
```

Average Play Time in last 2 we

| Game Names	
Football Manager 2016	1
Football Manager 2015	1531 m
Dota 2	1298 min
Stellaris	1282 min
Tree of Savior (English Ver.)	1217 min
ARK: Survival Evolved	1210 min
Clicker Heroes	1125 min
Total War: WARHAMMER	1086 min
FINAL FANTASY X/X-2 HD Remaster	1034 min
Factorio	961 min

```
# See if number of owners, total players, total player of 2 weeks, average play
time in last 2 weeks median total play time and yesterday's CUU correlated to e
ach other
library(corrplot)
```

```
## Warning: package 'corrplot' was built under R version 3.2.5
```

```
top100data_sub <- top100data[,6:11]
M <- cor(top100data_sub)
p3 <- corrplot(M, method="ellipse")
```
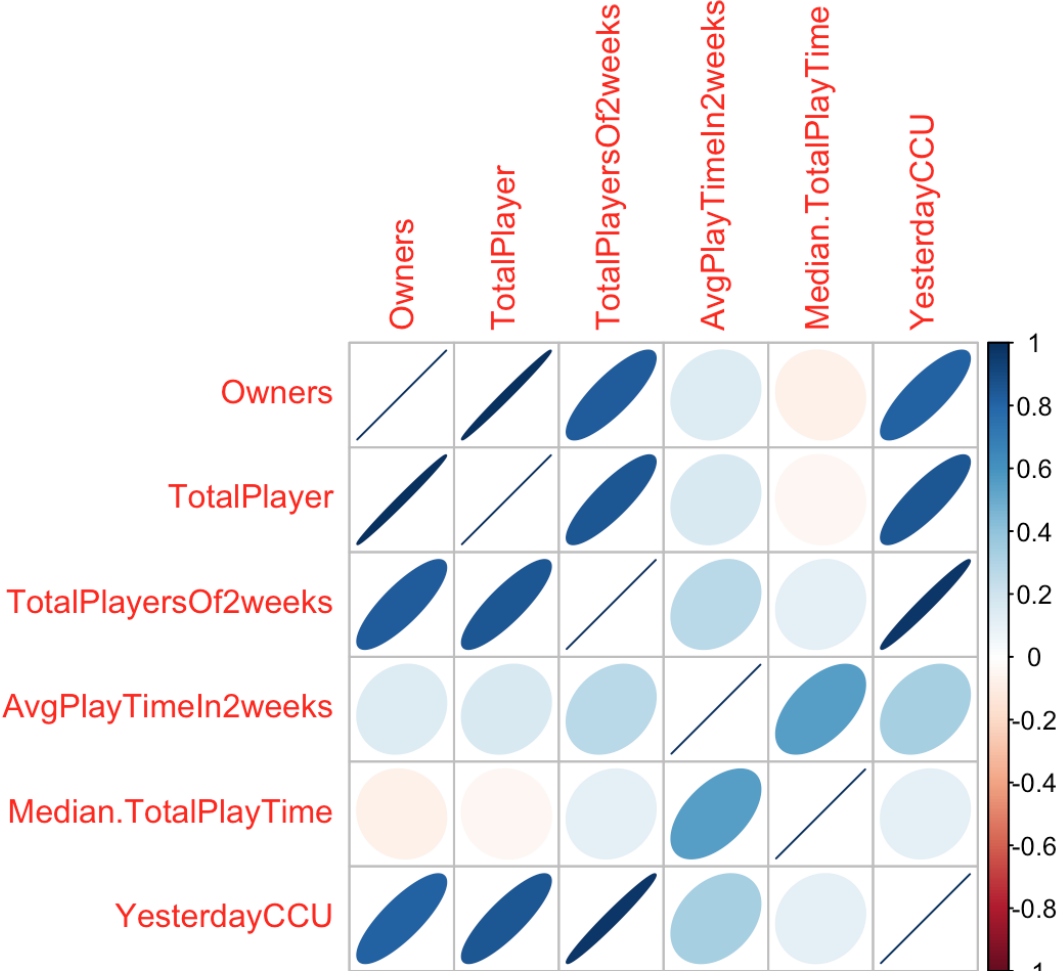
```
# It seems playing time of last 2 weeks doesn't corelate to total number of pla
yers
```