

Assignment 10

March 7, 2021

0.0.1 Portfolio assignment 10

15 min: Perform a bivariate analysis (Pearson correlation and scatter plot) on at least 1 combination of 2 columns with numeric data in the dataset that you chose in portfolio assignment 4. Does the correlation and scatter plot match your expectations? Add your answer to your notebook. Commit the Notebook to your portfolio when you're finished.

```
[7]: import pandas as pd
import seaborn as sns
```

```
[8]: steam = pd.read_csv('steam_games.csv')
steam.head()
```

```
[8]:
```

	url	types	\
0	https://store.steampowered.com/app/379720/DOOM/	app	
1	https://store.steampowered.com/app/578080/PLAY...	app	
2	https://store.steampowered.com/app/637090/BATT...	app	
3	https://store.steampowered.com/app/221100/DayZ/	app	
4	https://store.steampowered.com/app/8500/EVE_On...	app	

	name	\
0	DOOM	
1	PLAYERUNKNOWN'S BATTLEGROUNDS	
2	BATTLETECH	
3	DayZ	
4	EVE Online	

	desc_snippet	\
0	Now includes all three premium DLC packs (Unto...	
1	PLAYERUNKNOWN'S BATTLEGROUNDS is a battle roya...	
2	Take command of your own mercenary outfit of '...	
3	The post-soviet country of Chernarus is struck...	
4	EVE Online is a community-driven spaceship MMO...	

	recent_reviews	\
0	Very Positive,(554),- 89% of the 554 user revi...	
1	Mixed,(6,214),- 49% of the 6,214 user reviews ...	
2	Mixed,(166),- 54% of the 166 user reviews in t...	

3 Mixed,(932),- 57% of the 932 user reviews in t...
 4 Mixed,(287),- 54% of the 287 user reviews in t...

all_reviews release_date \
 0 Very Positive,(42,550),- 92% of the 42,550 use... May 12, 2016
 1 Mixed,(836,608),- 49% of the 836,608 user revi... Dec 21, 2017
 2 Mostly Positive,(7,030),- 71% of the 7,030 use... Apr 24, 2018
 3 Mixed,(167,115),- 61% of the 167,115 user revi... Dec 13, 2018
 4 Mostly Positive,(11,481),- 74% of the 11,481 u... May 6, 2003

developer publisher \
 0 id Software Bethesda Softworks,Bethesda Softworks
 1 PUBG Corporation PUBG Corporation,PUBG Corporation
 2 Harebrained Schemes Paradox Interactive,Paradox Interactive
 3 Bohemia Interactive Bohemia Interactive,Bohemia Interactive
 4 CCP CCP,CCP

popular_tags \
 0 FPS,Gore,Action,Demons,Shooter,First-Person,Gr...
 1 Survival,Shooter,Multiplayer,Battle Royale,PvP...
 2 Mechs,Strategy,Turn-Based,Turn-Based Tactics,S...
 3 Survival,Zombies,Open World,Multiplayer,PvP,Ma...
 4 Space,Massively Multiplayer,Sci-fi,Sandbox,MMO...

game_details \
 0 Single-player,Multi-player,Co-op,Steam Achieve...
 1 Multi-player,Online Multi-Player,Stats
 2 Single-player,Multi-player,Online Multi-Player...
 3 Multi-player,Online Multi-Player,Steam Worksho...
 4 Multi-player,Online Multi-Player,MMO,Co-op,Onl...

languages achievements \
 0 English,French,Italian,German,Spanish - Spain,... 54.0
 1 English,Korean,Simplified Chinese,French,Germa... 37.0
 2 English,French,German,Russian 128.0
 3 English,French,Italian,German,Spanish - Spain,... NaN
 4 English,German,Russian,French NaN

genre \
 0 Action
 1 Action,Adventure,Massively Multiplayer
 2 Action,Adventure,Strategy
 3 Action,Adventure,Massively Multiplayer
 4 Action,Free to Play,Massively Multiplayer,RPG,...

game_description \
 0 About This Game Developed by id software, the...

```

1 About This Game PLAYERUNKNOWN'S BATTLEGROUNDS...
2 About This Game From original BATTLETECH/Mech...
3 About This Game The post-soviet country of Ch...
4 About This Game

mature_content \
0 NaN
1 Mature Content Description The developers de...
2 NaN
3 NaN
4 NaN

minimum_requirements \
0 Minimum:,OS:,Windows 7/8.1/10 (64-bit versions...
1 Minimum:,Requires a 64-bit processor and opera...
2 Minimum:,Requires a 64-bit processor and opera...
3 Minimum:,OS:,Windows 7/8.1 64-bit,Processor:,I...
4 Minimum:,OS:,Windows 7,Processor:,Intel Dual C...

recommended_requirements original_price \
0 Recommended:,OS:,Windows 7/8.1/10 (64-bit vers... $19.99
1 Recommended:,Requires a 64-bit processor and o... $29.99
2 Recommended:,Requires a 64-bit processor and o... $39.99
3 Recommended:,OS:,Windows 10 64-bit,Processor:,... $44.99
4 Recommended:,OS:,Windows 10,Processor:,Intel i... Free

discount_price
0 $14.99
1 NaN
2 NaN
3 NaN
4 NaN

```

There's not a lot of numerical data in my dataset, in fact the only numerical data I have is achievements and price. I'm going to use a different dataset

```
[9]: vg = pd.read_csv('vgsales.csv')
vg.head()
```

```
[9]:
```

	Rank	Name	Platform	Year	Genre	Publisher	\
0	1	Wii Sports	Wii	2006.0	Sports	Nintendo	
1	2	Super Mario Bros.	NES	1985.0	Platform	Nintendo	
2	3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	
3	4	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	
4	5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	

	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
0	41.49	29.02	3.77	8.46	82.74

1	29.08	3.58	6.81	0.77	40.24
2	15.85	12.88	3.79	3.31	35.82
3	15.75	11.01	3.28	2.96	33.00
4	11.27	8.89	10.22	1.00	31.37

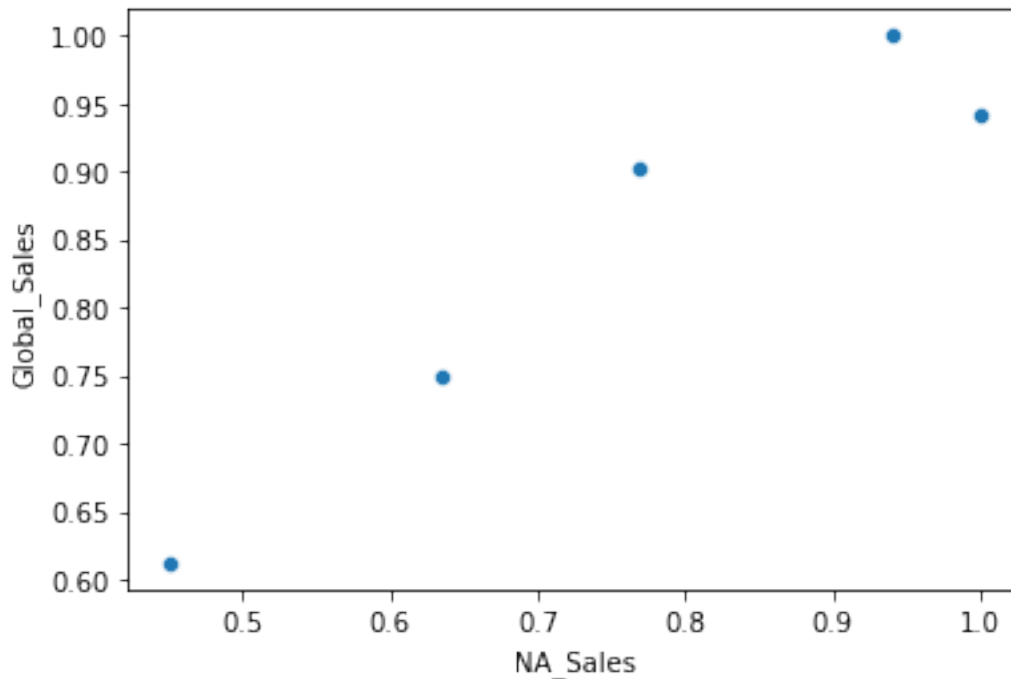
Let's check the bivariate correlation between countries.

```
[10]: corr = vg[['EU_Sales', 'NA_Sales', 'JP_Sales', 'Other_Sales', 'Global_Sales']].
      ↪corr()
      corr.style.background_gradient(cmap='coolwarm', axis=None).set_precision(2)
```

```
[10]: <pandas.io.formats.style.Styler at 0x252be99e730>
```

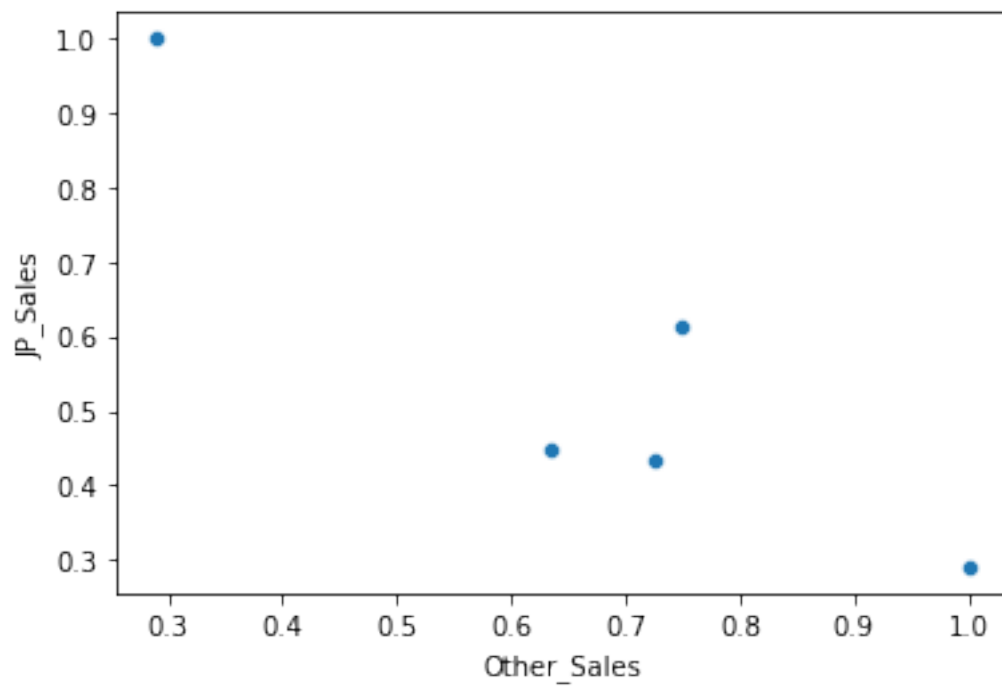
```
[11]: sns.scatterplot(x=corr.NA_Sales, y=corr.Global_Sales) # Strongest positive
```

```
[11]: <AxesSubplot:xlabel='NA_Sales', ylabel='Global_Sales'>
```



```
[12]: sns.scatterplot(x=corr.Other_Sales, y=corr.JP_Sales) # Weakest
```

```
[12]: <AxesSubplot:xlabel='Other_Sales', ylabel='JP_Sales'>
```



If I compare this to the penguin dataset, it does seem to be the same result.