

Business Research and Data Analytics

Lecture 6: MS Excel practicum: data storytelling

Igor Vyshnevskiy

Woosong University

April 10/13, 2024

Agenda

1. Groups allocation
2. Practicum: data storytelling
3. Results presentation

1. Groups allocation

Your groups (by ID)

G1	202112097	202010173	202312091	202310014	202312151	202312361
G2	202112157	202110085	202212224	202312035	202312163	202310003
G3	202212136	202110097	202310019	202312081	202312132	
G4	202212195	202310001	202312093	202312135	202212011	
G5	202212198	202310004	202001112	202310017	202201181	
G6	202212206	202310007	202201135	202312114	202312086	
G7	202212214	202310009	202201176	202312146	202312392	202312084

2. Practicum: data storytelling

Data

1. LMS Week 6 File: Practicum.csv

2. Data Source:

<https://archive.ics.uci.edu/dataset/468/online+shoppers+purchasing+intention+dataset>

What we do

- Imagine that your team is working on the e-commerce consulting project.
- You have collected some data and now want to get a valuable insight from that (i.e.,).
- Use a Practicum.csv file
- Explore your data
- Clean and transform if needed
- Build a dashboard (bunch of graphs and tables)
- Data Communication (i.e., *presentation*)
- Conclusions

File to be used: Practicum.csv

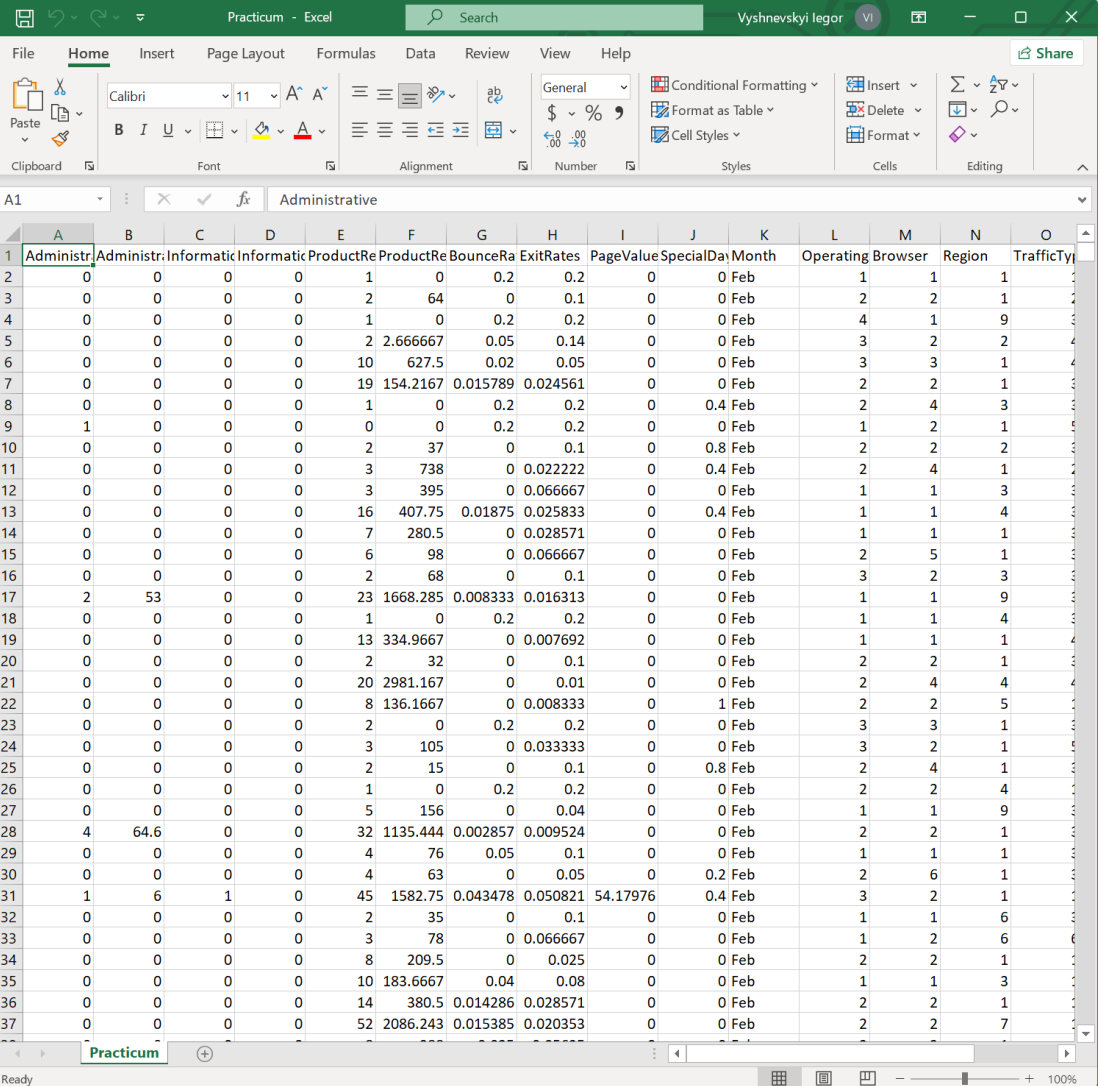
Our Data

Let's:

1. Observe our data. Is your data clean?;
2. Think on what to show/want to show.

The Online Shoppers Purchasing Intention dataset is a collection of data related to purchase patterns and consumer behaviour in the context of online shopping.

It was created by conducting surveys of online shoppers and collecting data from their responses.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Administrative	Administrative	Informatic	Informatic	Product	Product	BounceRate	ExitRates	PageValue	SpecialDay	Month	Operating	Browser	Region	TrafficType
2	0	0	0	0	1	0	0.2	0.2	0	0	Feb	1	1	1	1
3	0	0	0	0	2	64	0	0.1	0	0	Feb	2	2	1	2
4	0	0	0	0	1	0	0.2	0.2	0	0	Feb	4	1	9	2
5	0	0	0	0	2	2.666667	0.05	0.14	0	0	Feb	3	2	2	4
6	0	0	0	0	10	627.5	0.02	0.05	0	0	Feb	3	3	1	4
7	0	0	0	0	19	154.2167	0.015789	0.024561	0	0	Feb	2	2	1	4
8	0	0	0	0	1	0	0.2	0.2	0	0.4	Feb	2	4	3	2
9	1	0	0	0	0	0	0.2	0.2	0	0	Feb	1	2	1	5
10	0	0	0	0	2	37	0	0.1	0	0.8	Feb	2	2	2	2
11	0	0	0	0	3	738	0	0.022222	0	0.4	Feb	2	4	1	4
12	0	0	0	0	3	395	0	0.066667	0	0	Feb	1	1	3	3
13	0	0	0	0	16	407.75	0.01875	0.025833	0	0.4	Feb	1	1	4	3
14	0	0	0	0	7	280.5	0	0.028571	0	0	Feb	1	1	1	1
15	0	0	0	0	6	98	0	0.066667	0	0	Feb	2	5	1	3
16	0	0	0	0	2	68	0	0.1	0	0	Feb	3	2	3	3
17	2	53	0	0	23	1668.285	0.008333	0.016313	0	0	Feb	1	1	9	2
18	0	0	0	0	1	0	0.2	0.2	0	0	Feb	1	1	4	3
19	0	0	0	0	13	334.9667	0	0.007692	0	0	Feb	1	1	1	4
20	0	0	0	0	2	32	0	0.1	0	0	Feb	2	2	1	1
21	0	0	0	0	20	2981.167	0	0.01	0	0	Feb	2	4	4	4
22	0	0	0	0	8	136.1667	0	0.008333	0	1	Feb	2	2	5	2
23	0	0	0	0	2	0	0.2	0.2	0	0	Feb	3	3	1	1
24	0	0	0	0	3	105	0	0.033333	0	0	Feb	3	2	1	5
25	0	0	0	0	2	15	0	0.1	0	0.8	Feb	2	4	1	1
26	0	0	0	0	1	0	0.2	0.2	0	0	Feb	2	2	4	1
27	0	0	0	0	5	156	0	0.04	0	0	Feb	1	1	9	6
28	4	64.6	0	0	32	1135.444	0.002857	0.009524	0	0	Feb	2	2	1	3
29	0	0	0	0	4	76	0.05	0.1	0	0	Feb	1	1	1	1
30	0	0	0	0	4	63	0	0.05	0	0.2	Feb	2	6	1	3
31	1	6	1	0	45	1582.75	0.043478	0.050821	54.17976	0.4	Feb	3	2	1	1
32	0	0	0	0	2	35	0	0.1	0	0	Feb	1	1	6	2
33	0	0	0	0	3	78	0	0.066667	0	0	Feb	1	2	6	6
34	0	0	0	0	8	209.5	0	0.025	0	0	Feb	2	2	1	1
35	0	0	0	0	10	183.6667	0.04	0.08	0	0	Feb	1	1	3	1
36	0	0	0	0	14	380.5	0.014286	0.028571	0	0	Feb	2	2	1	1
37	0	0	0	0	52	2086.243	0.015385	0.020353	0	0	Feb	2	2	7	1

Our Data (cont.)

Some of the variables in this dataset include:

1. Administrative - The number of pages of the website visited by the user for administrative purposes
2. Administrative_Duration - The total time spent by the user on administrative pages of the website
3. Informational - The number of pages of the website visited by the user for informational purposes
4. Informational_Duration - The total time spent by the user on informational pages of the website
5. ProductRelated - The number of pages of the website visited by the user for product-related purposes
6. ProductRelated_Duration - The total time spent by the user on product-related pages of the website
7. BounceRates - The percentage of visitors who enter the website and leave without viewing any other pages
8. ExitRates - The percentage of visitors who exit the website from a particular page after visiting it
9. PageValues - The average value of the pages viewed by the user before the transaction
10. SpecialDay - The proximity of the visit to a special day (e.g., Mother's Day, Valentine's Day, etc.)

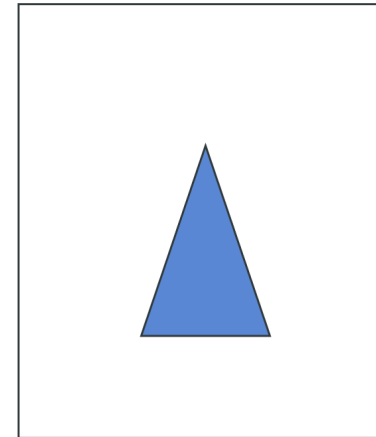
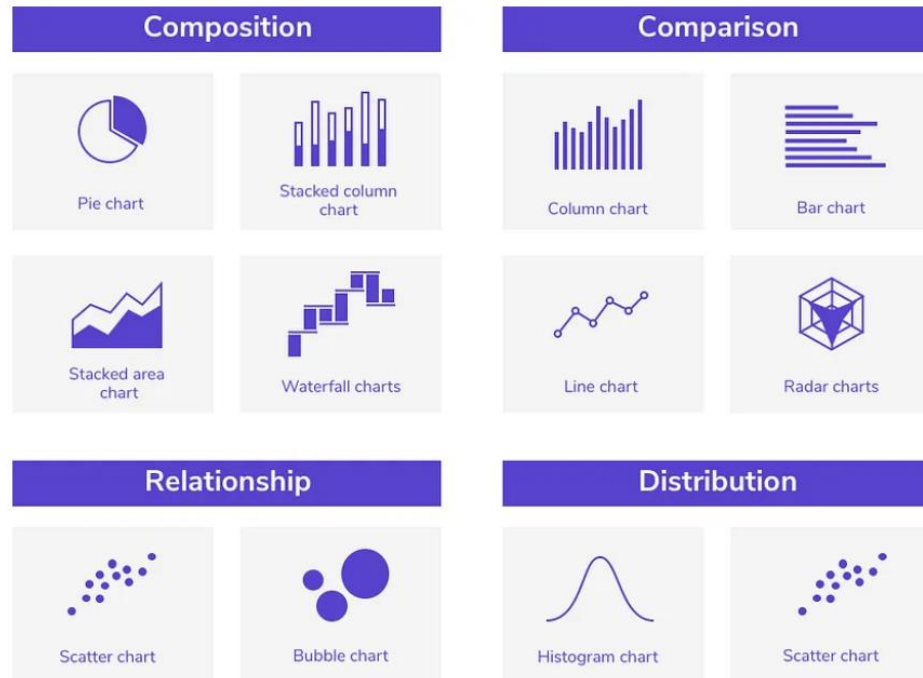
Our Data (cont.)

Additional Variable Information:

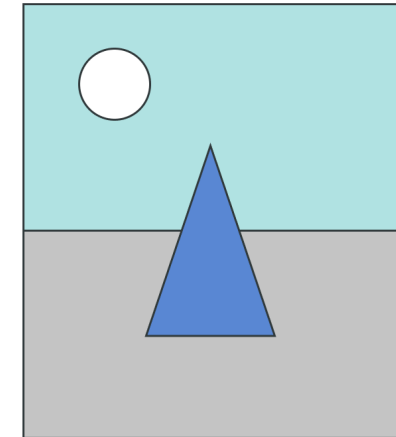
1. "Administrative", "Administrative Duration", "Informational", "Informational Duration", "Product Related" and "Product Related Duration" represent the number of different types of pages visited by the visitor in that session and total time spent in each of these page categories. The values of these features are derived from the URL information of the pages visited by the user and updated in real time when a user takes an action, e.g. moving from one page to another.
2. The "Bounce Rate", "Exit Rate" and "Page Value" features represent the metrics measured by "Google Analytics" for each page in the e-commerce site.
 1. The value of "Bounce Rate" feature for a web page refers to the percentage of visitors who enter the site from that page and then leave ("bounce") without triggering any other requests to the analytics server during that session.
 2. The value of "Exit Rate" feature for a specific web page is calculated as for all pageviews to the page, the percentage that were the last in the session.
 3. The "Page Value" feature represents the average value for a web page that a user visited before completing an e-commerce transaction.
3. The "Special Day" feature indicates the closeness of the site visiting time to a specific special day (e.g. Mother's Day, Valentine's Day) in which the sessions are more likely to be finalized with transaction. The value of this attribute is determined by considering the dynamics of e-commerce such as the duration between the order date and delivery date.
 1. For example, for Valentine's day, this value takes a nonzero value between February 2 and February 12, zero before and after this date unless it is close to another special day, and its maximum value of 1 on February 8.
4. The dataset also includes operating system, browser, region, traffic type, visitor type as returning or new visitor, a Boolean value (TRUE or FALSE) indicating whether the date of the visit is weekend, and month of the year.
5. Revenue contains either a True or False value, which correspond to whether or not the user made a purchase on the website during their visit, respectively.

Story

- What can you tell from this data?
- What elements (tables / graphs) do you want to use?



Data



+ Storytelling

Your Story time

Finally, you can move to create your story.

Be ready to show / present your work by 2:50 pm.

Rembert to upload your work.

Technical notes:

- a. Choose a person (say, a team leader) who will do the uploading and reporting of your group work.
- b. Remember to write down the names and IDs of all actual participants.
- c. The team leader briefly reports on the group members' participation activity and each member contribution (to avoid free riding).
- d. Upload the file (PPT) on time. Good luck!