



Data Analyst Case

For Candidate

Times

3 days to doing a Test Case

Instruction

An owner of a company wants to make the company's marketing strategy more effective. His strategy is to categorize each customer into a certain group so that he can sell the products to the relevant customers only. He also requires that each customer belongs to one group only.

Here is the list of attributes:

CustID = customer ID (unique)

Gender = gender of the customer

Age = age of the customer

AnnualIncome = annual income of the customer

CustScore = a score assigned by the store based on customer's purchase behavior (1-100).

Score '1' means the intensity of purchase is at the lowest level. Score '100' means the intensity

of purchase is at the highest level

Hometown = hometown of the customer

FavoriteTime = favorite time for shopping

- a. From the provided data (customer_segmentation_data.csv), try to do Exploratory Data Analysis. Expected output should be some visualization of graph or chart with the explanation.
- b. Try to give a suggestion how to do the segmentation based on the A result, give us about the reason behind that. (Ex: using one of classification/clustering method of machine learning)

Table Invoices

No	company_name	invoice_id	date
1	Sarale Store	e588bdd5-29f2-4caa-b361-82e4f63e781b	2017-09-30
2	wira chat	3a0de6ee-8c05-4008-8074-7dc85d87da21	2018-03-08
3	wira chat	9ccacdd4-5eed-4ad0-9883-65aee1b6e7ee	2018-03-08
4	wira chat	cf74b6cc-83ae-4d60-9349-714307428a07	2018-03-12
5	wira chat	a9c6658f-fea1-4d6e-abd4-a9b76faadeca	2018-03-26
6	wira chat	750f20ea-12c4-40c9-b834-e1c8e5686c25	2018-03-26
7	wira chat	a9370794-112a-4960-b4cd-caf4c3d61cf6	2018-03-26
8	wira chat	eb745ce2-dae5-4a2e-8d34-ba9e37caba8d	2018-03-26
9	wira chat	17574eea-2a84-42e3-bc25-77e8f9509383	2018-03-28
10	wira chat	bba534b9-082e-4c3a-b78a-21d01c0a6ee4	2018-03-31
11	wira chat	b5cdd61e-23b4-4421-b5bc-5a5c2c6e23af	2018-03-31
12	Penjualan handphone	1f4e91bb-06d6-4657-b882-a612c3026ccc	2018-03-27
13	PT CAKRAWALA INDONESIA SEJAHTERA	4d68adee-e835-4660-a56f-8579527fb06d	2018-07-01
14	Tometo Club	020d7ffc-492d-4cac-802f-32b1b34ddc60	2018-02-05
15	Kantor Notaris/PPAT Viondi Yunatan, S.H., M.Kn.	7e63ddcd-f25b-4d31-aae4-cfc93c54289b	2018-03-25
16	Sparepart Laptop	61e58eaf-68f2-49bf-a9d5-d6b3a487ab6d	2019-04-11
17	PT PERMATA SOLUSI INDONESIA	f245ff43-db3b-42d0-a207-d78822fe3f79	2018-06-05
18	PT PERMATA SOLUSI INDONESIA	87f94d55-7e9a-41b1-90b6-98f038c8354c	2018-06-05
19	ELISABET HOUSE	4fa0e9f0-00a9-40a9-8400-f07706154613	2018-03-26
20	PT Lindung Karya Indah	e2f940f1-37ea-4042-81e9-f6ddbd974198	2016-10-26

Table Invoice Items

No	company_name	invoice_id	date	item_name	price
1	Sarale Store	e588bdd5-29f2-4caa-b361-82e4f63e781b	2017-09-30	Ongkir1	11000
2	Sarale Store	e588bdd5-29f2-4caa-b361-82e4f63e781b	2017-09-30	Packing	1000
3	Sarale Store	e588bdd5-29f2-4caa-b361-82e4f63e781b	2017-09-30	T-shirt Long Sleeve	75000
4	Sarale Store	e588bdd5-29f2-4caa-b361-82e4f63e781b	2017-09-30	T-shirt Short Sleeve	70000
5	wira chat	b5cdd61e-23b4-4421-b5bc-5a5c2c6e23af	2018-03-31	001 - Plugins WP	200000
6	wira chat	17574eea-2a84-42e3-bc25-77e8f9509383	2018-03-28	001 - Plugins WP	200000
7	wira chat	750f20ea-12c4-40c9-b834-e1c8e5686c25	2018-03-26	001 - Plugins WP	200000
8	wira chat	9ccacdd4-5eed-4ad0-9883-65aee1b6e7ee	2018-03-08	001 - Plugins WP	200000
9	wira chat	a9c6658f-fea1-4d6e-abd4-a9b76faadeca	2018-03-26	001 - Plugins WP	200000
10	wira chat	cf74b6cc-83ae-4d60-9349-714307428a07	2018-03-12	001 - Plugins WP	200000
11	wira chat	3a0de6ee-8c05-4008-8074-7dc85d87da21	2018-03-08	001 - Plugins WP	200000
12	wira chat	bba534b9-082e-4c3a-b78a-21d01c0a6ee4	2018-03-31	001 - Plugins WP	200000
13	wira chat	eb745ce2-dae5-4a2e-8d34-ba9e37caba8d	2018-03-26	001 - Plugins WP	200000
14	wira chat	a9370794-112a-4960-b4cd-caf4c3d61cf6	2018-03-26	001 - Plugins WP	200000
15	Penjualan handphone	1f4e91bb-06d6-4657-b882-a612c3026ccc	2018-03-27	xiaomi	1000000
16	PT CAKRAWALA INDONESIA SEJAHTERA	4d68adee-e835-4660-a56f-8579527fb06d	2018-07-01	Sign On	70000
17	Tometo Club	020d7ffc-492d-4cac-802f-32b1b34ddc60	2018-02-05	Company Profile	3500000
18	Kantor Notaris/PPAT Viondi Yunatan, S.H., M.Kn.	7e63ddcd-f25b-4d31-aae4-cfc93c54289b	2018-03-25	PNBP	1680000
19	Kantor Notaris/PPAT Viondi Yunatan, S.H., M.Kn.	7e63ddcd-f25b-4d31-aae4-cfc93c54289b	2018-03-25	Akta Notaris	4000000
20	Sparepart Laptop	61e58eaf-68f2-49bf-a9d5-d6b3a487ab6d	2019-04-11	RAM Laptop	300000
21	PT PERMATA SOLUSI INDONESIA	87f94d55-7e9a-41b1-90b6-98f038c8354c	2018-06-05	AMS 2018	9604000
22	PT PERMATA SOLUSI INDONESIA	f245ff43-db3b-42d0-a207-d78822fe3f79	2018-06-05	CloverHR	38082800
23	ELISABET HOUSE	4fa0e9f0-00a9-40a9-8400-f07706154613	2018-03-26	Dormitory	1100000
24	ELISABET HOUSE	4fa0e9f0-00a9-40a9-8400-f07706154613	2018-03-26	Dormitory	1200000
25	ELISABET HOUSE	4fa0e9f0-00a9-40a9-8400-f07706154613	2018-03-26	Dormitory	1600000
26	ELISABET HOUSE	4fa0e9f0-00a9-40a9-8400-f07706154613	2018-03-26	Dormitory	1600000
27	ELISABET HOUSE	4fa0e9f0-00a9-40a9-8400-f07706154613	2018-03-26	Dormitory	2000000
28	ELISABET HOUSE	4fa0e9f0-00a9-40a9-8400-f07706154613	2018-03-26	Rent 1 House	7000000
29	ELISABET HOUSE	4fa0e9f0-00a9-40a9-8400-f07706154613	2018-03-26	Rent 1 House	1500000
30	ELISABET HOUSE	4fa0e9f0-00a9-40a9-8400-f07706154613	2018-03-26	Rent 1 House	6800000
31	PT Lindung Karya Indah	e2f940f1-37ea-4042-81e9-f6ddbd974198	2016-10-26	Shoes	50000

Above are the tables for invoices. The first table (Table Invoices) is a table for showing every invoice that has been recorded, and the second table (Table Invoice Items) is a table for showing every item in each invoice. Based on this, try to make a query to answer those questions below:

- a. Number of invoices that every company makes on each date
- b. Number of items for every invoice for each company
- c. Price for every invoice for each company

Specification and limitations

- **For Question Number 1:** You are free to use any coding language such as Python or R & library that you are familiar with, but make sure you understand it pretty well.
- Expected output is a report with the details of the result. You have to be as detailed as possible such that we can reproduce your work but also as concise as possible (e.g. prefer bullets over unnecessary lengthy paragraphs)
- **For Question Number 2:** Solve it using a written SQL. Write your query directly on the report that you create.