

Exploratory Data Analysis

G2M insight for Cab Investment firm

17/07/2022

Executive Summary

- In the following presentation we will to analyze two companies in the Cab industry, in order to get a better understanding of the market and make the right decision as to which company XYZ should invest in.

Problem Statement

- In our analysis we relied on 4 data files:
 1. Cab_Data.csv: Detailed the transactions (Cost/Revenue/Km traveled...)
 2. City.csv: Gave the number of users by city
 3. Customer_ID.csv: Information about each user (Gender/Age/Income)
 4. Transaction_ID.csv: Connected between the transaction and the user.

- As we are dealing with 4 separate files, it is hard to get as much information as if they were all combined in one file. We had to first get a table that groups the data.

Approach

- Using Power Query in Excel, we are able to connect between customer ID (in file Customer_ID.csv) and transaction ID (in file Transaction_ID.csv) and file Cab_Data.csv:

Clipboard			Font			Alignment			Number			Styles		
O2														
Transaction ID	Date of Travel	Company	City	KM Travelled	Price Charged	Cost of Trip	Profit	Customer ID	Payment_Mode	Gender	Age	Income (USD/Month)		
10000013	02/01/2016 00:00	Pink Cab	ATLANTA GA	9,04	125,2	97,632	27,568	28712	Cash	Male	53	11242		
10012732	04/03/2016 00:00	Pink Cab	NEW YORK NY	42,4	703,09	457,92	245,17	1	Card	Male	36	16359		
10000041	02/01/2016 00:00	Pink Cab	CHICAGO IL	35,02	598,43	406,232	192,198	4289	Card	Male	19	28719		
10000045	02/01/2016 00:00	Pink Cab	CHICAGO IL	3,24	48,04	33,696	14,344	5734	Card	Male	23	22475		
10000048	02/01/2016 00:00	Pink Cab	CHICAGO IL	14,04	223,6	140,4	83,2	5915	Card	Male	26	24115		
10000049	02/01/2016 00:00	Pink Cab	CHICAGO IL	27,82	451,22	303,238	147,982	5628	Cash	Male	33	15231		
10000050	02/01/2016 00:00	Pink Cab	CHICAGO IL	10,2	165,77	116,28	49,49	3125	Cash	Male	25	9562		
10000055	02/01/2016 00:00	Pink Cab	CHICAGO IL	16,95	249,42	200,01	49,41	3600	Cash	Male	20	4277		
10000060	02/01/2016 00:00	Pink Cab	CHICAGO IL	35,34	500,22	395,808	104,412	3330	Card	Male	46	5216		
10000066	02/01/2016 00:00	Pink Cab	DALLAS TX	13,13	122,42	149,682	-27,262	25155	Card	Male	27	10997		
10000067	02/01/2016 00:00	Pink Cab	DALLAS TX	33,32	308,58	386,512	-77,932	25247	Cash	Male	26	24178		
10000070	02/01/2016 00:00	Pink Cab	DENVER CO	7,02	61,3	82,836	-21,536	30718	Cash	Male	52	20255		
10226027	17/09/2017 00:00	Pink Cab	NEW YORK NY	23,92	468,05	282,256	185,794	1	Cash	Male	36	16359		
10000074	02/01/2016 00:00	Pink Cab	LOS ANGELES CA	46	997,84	542,8	455,04	8693	Card	Male	21	15024		
10000092	02/01/2016 00:00	Pink Cab	LOS ANGELES CA	37,76	851,25	438,016	413,234	8927	Card	Male	19	17197		
10000113	02/01/2016 00:00	Pink Cab	LOS ANGELES CA	12,32	265,25	137,984	127,266	7924	Cash	Male	26	23137		
10241365	21/10/2017 00:00	Pink Cab	NEW YORK NY	38,5	733,87	462	271,87	1	Cash	Male	36	16359		
10000119	02/01/2016 00:00	Pink Cab	LOS ANGELES CA	22,31	463,5	243,179	220,321	7199	Card	Male	31	20802		
10000124	02/01/2016 00:00	Pink Cab	LOS ANGELES CA	37,1	803,58	408,1	395,48	6465	Card	Male	46	3249		
10000127	02/01/2016 00:00	Pink Cab	LOS ANGELES CA	38	850,87	425,6	425,27	7539	Card	Male	40	24169		
10000131	02/01/2016 00:00	Pink Cab	LOS ANGELES CA	12	251,46	122,4	129,06	8751	Cash	Male	34	17745		
10000135	02/01/2016 00:00	Pink Cab	MIAMI FL	6,42	105,42	64,842	40,578	9584	Card	Male	31	22812		
10270727	10/12/2017 00:00	Pink Cab	NEW YORK NY	20,7	381,47	211,14	170,33	1	Card	Male	36	16359		
10000145	02/01/2016 00:00	Pink Cab	NEW YORK NY	2,1	37,18	21,42	15,76	502	Cash	Male	28	15285		
10000149	02/01/2016 00:00	Pink Cab	NEW YORK NY	32,64	498,6	349,248	149,352	533	Card	Male	52	15974		
10000163	02/01/2016 00:00	Pink Cab	PHOENIX AZ	4,44	71,57	48,84	22,73	22557	Cash	Male	38	8808		
10000164	02/01/2016 00:00	Pink Cab	PHOENIX AZ	8,55	114,15	89,775	24,375	22469	Card	Male	37	4378		
10000171	02/01/2016 00:00	Pink Cab	SAN DIEGO CA	14,28	269,15	147,084	122,066	20687	Cash	Male	39	8926		
10000182	02/01/2016 00:00	Pink Cab	SAN DIEGO CA	2,36	47,89	23,6	24,29	18418	Cash	Male	39	13112		
10000187	02/01/2016 00:00	Pink Cab	SAN DIEGO CA	23,8	485,21	252,28	232,93	18817	Cash	Male	40	5038		
10000189	02/01/2016 00:00	Pink Cab	SAN DIEGO CA	41,76	793,63	476,064	317,566	18488	Cash	Male	37	22441		
10000195	02/01/2016 00:00	Pink Cab	SAN DIEGO CA	32,3	654,5	368,22	286,28	20752	Cash	Male	63	18683		
10000201	02/01/2016 00:00	Pink Cab	SAN DIEGO CA	31,68	623,77	370,656	253,114	18490	Card	Male	24	10573		
10000345	02/01/2016 00:00	Pink Cab	SEATTLE WA	20,4	259,48	234,6	24,88	37896	Card	Male	39	7143		
10000349	02/01/2016 00:00	Pink Cab	SEATTLE WA	7,21	90,93	72,1	18,83	36691	Cash	Male	28	6491		
10000351	02/01/2016 00:00	Pink Cab	SEATTLE WA	2,08	24,79	21,008	3,782	38641	Card	Male	24	13801		
10000356	02/01/2016 00:00	Pink Cab	SEATTLE WA	14,95	171,47	171,925	-0,455	38635	Card	Male	31	24682		

EDA

- DATA EXPLORATION:

We have a total of 355,032 data points

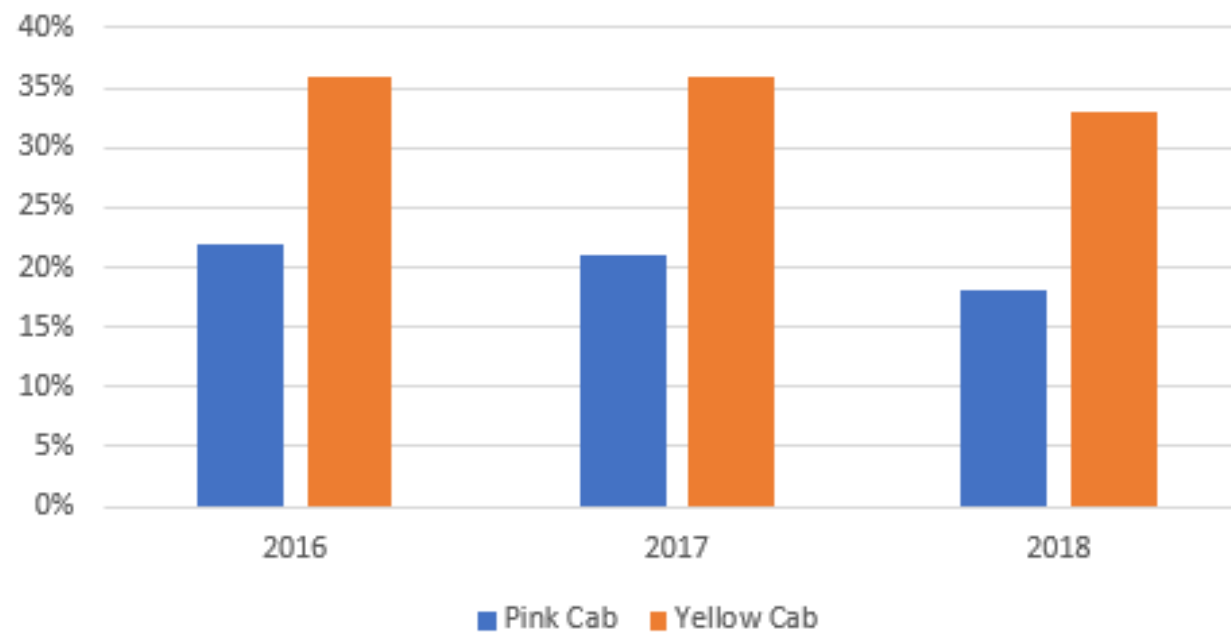
We also have 4 .csv files:

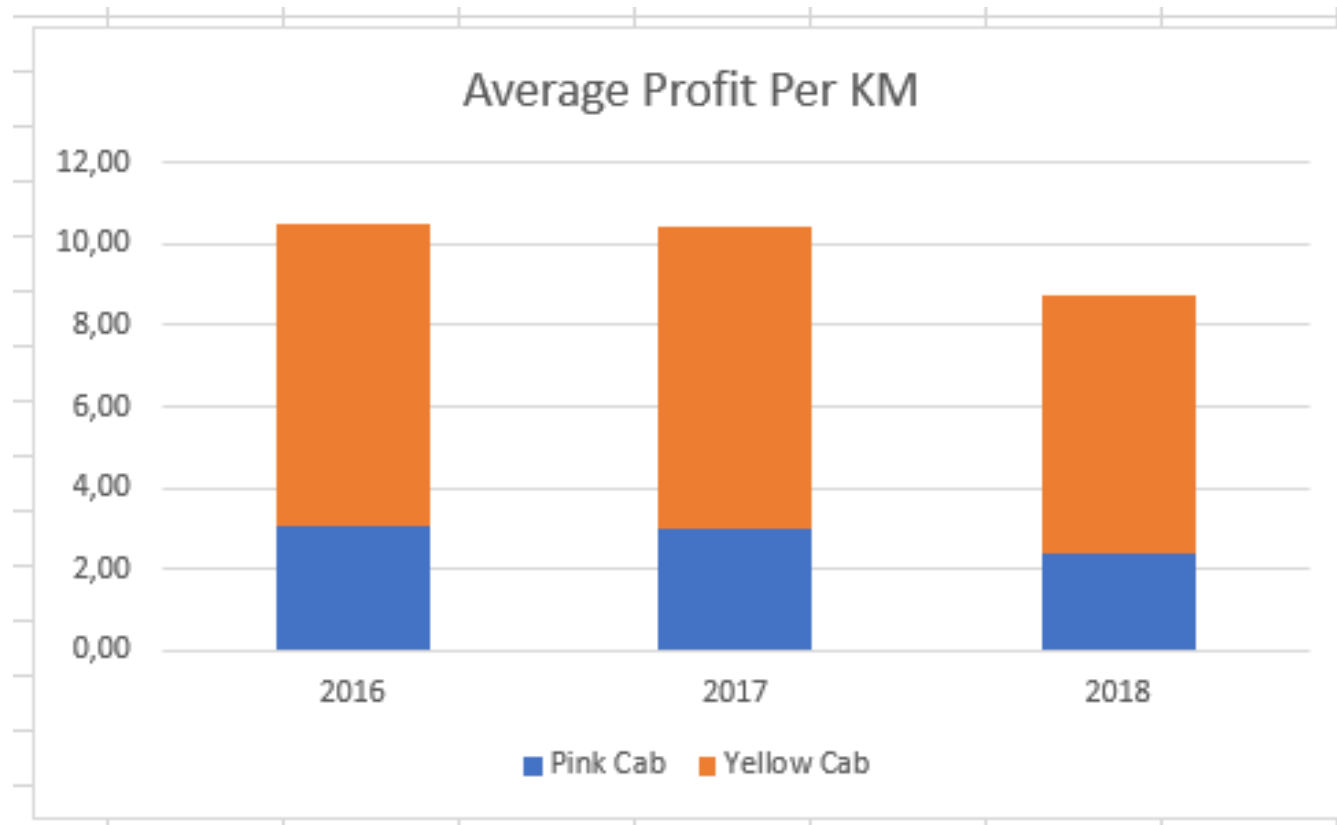
- Cab_Data.csv: 7 features
- City.csv: 3 features
- Customer_ID.csv: 4 features
- Transaction_ID.csv: 3 features

Profit Analysis

Company ▼	Total Price Charged ▼	Total Costs ▼	Profit ▼
Pink Cab	\$26.328.251,33	\$21.020.923,01	\$5.307.328,32
Yellow Cab	\$125.853.887,20	\$81.833.514,02	\$44.020.373,17

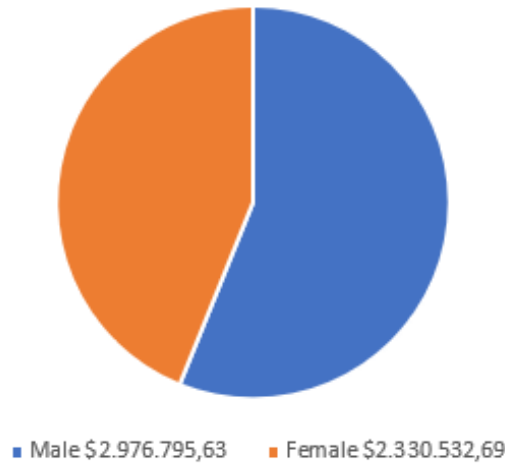
Profit by Year



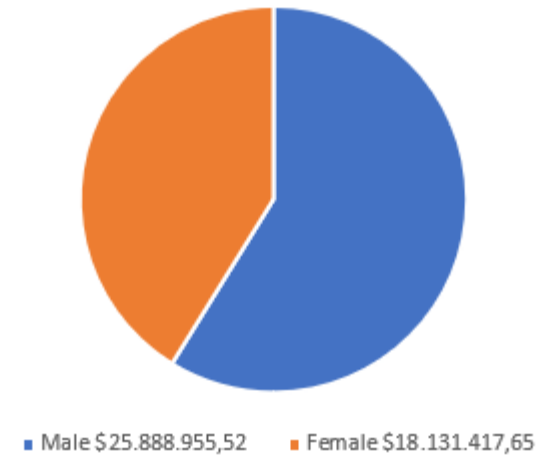


Profit Gender Wise

PINK CAB: Profit Contribution Gender Wise

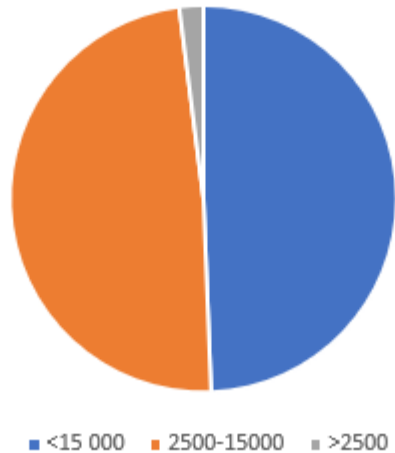


YELLOW CAB: Profit Per

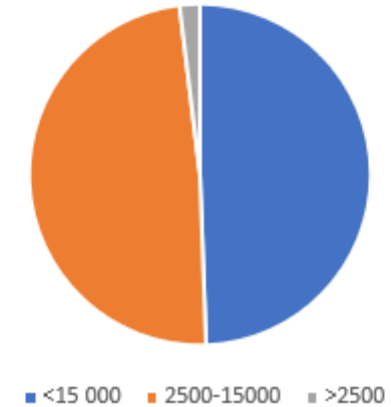


Profit Based on Income of Customers

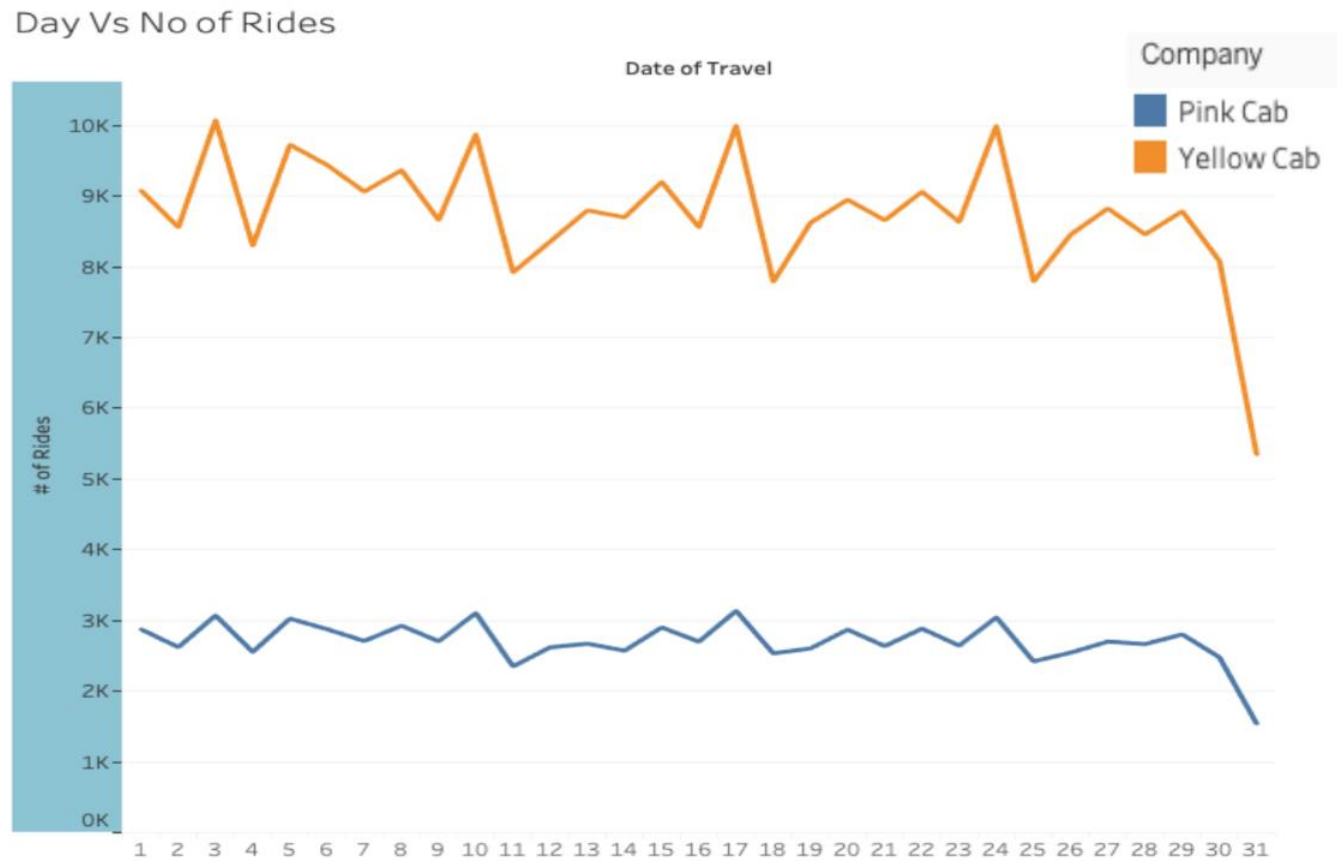
PINK CAB: Profit based on income of customers



YELLOW CAB: Profit based on income of customers



Seasonality of demand



RECOMMENDATIONS

- We have evaluated both companies on the following points:
 - I. Profit of Previous 3 years: Yellow cab has made around 9 times more profit than Pink Cab did during the period of 2016-2018
 - II. Profit by Year: We observed a decline in profit in both companies, nevertheless, Yellow Cab still has a much higher profit percentage each year averaging around 34.6% profit, while Pink Cab only averages 18.97%.
 - III. Average Profit Per KM: Yellow Cab profits almost 2.6 more times than Pink Cab does in each km traveled
 - IV. Profit Gender Wise: Males contribute the most to both companies profits and at around the same percentage.
 - V. Profit based on income: By analyzing the data, we observed that customers, in both companies, belong to similar social class. We also notice that both companies make the most profit off of customers with higher income (>15 000).
 - VI. Seasonality in the demand: We noticed that the demand spikes in both companies on the weekends. However, the demand on Yellow cab services remain significantly higher.

Based on those points, we recommend the investment in Yellow Cab.