

PROBLEM STATEMENT: Customers prefer the Ola app to reach the railway station but prefer other cab services or modes of transport while leaving the station.

ASSUMPTIONS MADE FOR THIS CASE STUDY :

- Consider a year with no Economic Recession
- Regular month (no holiday/festive season)
- No sudden hike in petrol/cab/airfare
- Location- Mumbai and the airport is located in the city
- No political unrest - no public/private transportation strike
- Regular traffic
- Day- weekday
- Equal fare prices for all cab companies eg: Uber and Ola
- Equal waiting period (ETA) for all cab companies.

REASONS WHY CUSTOMER CHOOSES OLA FOR DROP-OFFS :

1. The customer wants a reliable mode of transport to the railway station as they do not want to miss their train. Therefore, they prefer Time and Reliability over Cost.
2. The customer books an Ola cab in the comfort of their home where there is no form of marketing strategy being used to promote other cab companies(eg: Uber) or other modes of transport (eg: Traditional Taxi, Bus, Autos). Therefore, they have already made up their mind based on their previous experience of Ola.
3. The customer does not have to leave their home to hail a cab thus saving Time and Effort.
4. The customer chooses Ola over other modes of transport (for safety reasons) when the train departure is scheduled late at night or very early in the morning.
5. The customer has the flexibility to cancel their ride if the trains are delayed or cancelled.

REASONS WHY CUSTOMERS DID NOT CHOOSE OLA FOR PICKUPS :

1. When the customer arrives from their trip, they are less worried and more relaxed about reaching home quickly and hence would not mind finding cheaper options than Ola. Therefore, they prefer Cost over Time and Reliability.
2. The customer might choose Local Trains as their mode of transport as they are already at the station, and would be much cheaper and faster than Ola.

3. At the Station, the customer is a target for a lot of strategically placed Ads which can cause them to change their preference from Ola.
4. Some customers like to negotiate/ haggle, and since it is not possible w Ola/Uber, many might prefer Traditional Taxis as Time and Reliability is not an issue anymore.
5. Family/friends are more likely to pick up than drop off.
6. Poor internet connection

SUGGESTIONS :

1. Offer the customer a Round-Trip Payment Plan where he/she pays for drop-offs and can pre-book pickups at the same time at a discounted price.
2. Use the Railway's API to determine when the trains arrive in real-time as they tend to get delayed a lot. This feature can be used by the drivers to arrive at the station to equate supply with demand.
3. Offer ride-sharing options as the pickup for all passengers will be the same.
4. Offer customers some discount on entering their travel itinerary details - day, date and timing of their arrival and then sending them a message/notification an hour before they arrive at the station. Also, an option to pre-book an hour before arrival to avoid waiting.

Q. What should be the Success Metric of Content Creation, Content Consumption and Monetization on the Platform Level?

ASSUMPTIONS: The following success metrics would be in regards to the Nojoto App which is the country's leading storytelling platform.

Content Creation :

1. Profile Clicks: It measures the number of people who liked the creator's video enough to check out their profile for similar content.
2. Followers: This indicates that the creator is promising and serves content to a wide range of demography and is good at posting regularly.
3. Trends or Original: This metric can be used to see if people re-create more content because of existing trends or are they more inclined towards creating their own original content. This can be measured using Hashtags.
4. Category Expansion: Average number of categories/languages a user makes his/her content in.

Content Consumption :

1. Number of Shares: The biggest metric for content consumption is when the content lands onto other social media apps (eg WhatsApp) which usually means the content is not only liked and consumed by Nojoto's user base but also the user base of other apps. Which in future might turn into potential users.
2. Total Number of Views: This number should steadily increase as the user base of the app increases.
3. Likes and Comments: This gives us a clear and quick view of what type of content is being consumed the most.
4. Average Play Time: The average playtime should be more than the time taken by the user to scroll. Therefore, Average playtime > Average Scroll time
5. Trending Category: Used to determine which category is always trending and has the most content, views, engagement and shares.
6. Demography: Used to determine which demography (eg: Age, Gender, Location) is most active on the app thus contributing to the engagement numbers.
7. Churn Rate: This metric is used to see how many users have deleted their account or have a ghost account within a particular time span.

8. Return Rate: This metric determines the number of people who install Nojoto and come back to use it on a particular day, week, or month.

Monetisation :

1. New Installs: Number of newly acquired users per month.
2. User Lifespan / Retention: This shows us the average user retention rate.
3. User Lifetime Value: This metric can be used to indicate the total revenue the app can expect from the single user account and for how long.
4. Media Mentions: This metric can be used to determine which other social media platform mentions Nojoto and/or harbours Nojoto's content.
5. Referrals and reviews

Q. What would you improve about our Nojoto App as Product & Why?

- IOS version
- Subtitles to increase accessibility.
- Option for only-text video content.

Q. Kindly Pick any dummy dataset and make a Cohort / Pivot using any Tool/Programming Language, and then Send the final output.

[GitHub Link](#), My assignment can be viewed here!

```
In [ ]: # This Python 3 environment comes with many helpful analytics libraries installed
# It is defined by the kaggle/python Docker image: https://github.com/kaggle/docker-python
# For example, here's several helpful packages to load

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

# Input data files are available in the read-only "../input/" directory
# For example, running this (by clicking run or pressing Shift+Enter) will list all files under the input directory

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))

# You can write up to 20GB to the current directory (/kaggle/working/) that gets preserved as output when you create a version
using "Save & Run All"
# You can also write temporary files to /kaggle/temp/, but they won't be saved outside of the current session
```

```
In [2]: import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        print(os.path.join(dirname, filename))
```

```
/kaggle/input/titanic/train.csv
/kaggle/input/titanic/test.csv
/kaggle/input/titanic/gender_submission.csv
```

```
In [3]: #import necessary libraries
import pandas as pd
import numpy as np
```

```
In [4]: df = pd.read_csv('../input/titanic/train.csv')
df.head()
```

```
Out[4]:
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cummings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

```
In [5]: #pivot using single index
df2 = pd.pivot_table(df, index=['Sex'])
df2
```

```
Out[5]:
```

	Age	Fare	Parch	PassengerId	Pclass	SibSp	Survived
Sex							
female	27.915709	44.479818	0.649682	431.028662	2.159236	0.694268	0.742038
male	30.726645	25.523893	0.235702	454.147314	2.389948	0.429809	0.188908

```
In [12]: #We can determine the possibility of a passenger surviving on the basis of their Gender and their Passenger Class.
df6 = pd.pivot_table(df, index=['Sex'], aggfunc={'Survived': np.sum})
df6
```

```
Out[12]:
```

	Survived
Sex	
female	233
male	109

OBSERVATION: Female significantly survived more than men

EDUCATED GUESS: During a tragedy, women and children are given more preference over men

```
In [10]: df5 = pd.pivot_table(df, index=['Sex', 'Pclass'], aggfunc={'Survived': np.sum})
df5
```

```
Out[10]:
```

		Survived
Sex	Pclass	
female	1	91
	2	70
	3	72
male	1	45
	2	17
	3	47

	S	12
male	1	45
	2	17
	3	47

OBSERVATION: Women in the first class had higher chance of survival.

EDUCATED GUESS: First class Women passengers were given preference (or alerted before) to board lifeboats and thus survived.

In [9]: *#We can determine the possibility of a passenger surviving on the basis of their Gender and the port they embarked from.*
#C = Cherbourg
#Q = Queenstown
#S = Southampton
df5 = pd.pivot_table(df,index=['Sex' , 'Embarked'],aggfunc={ 'Survived':np.sum})
df5

Out[9]:

		Survived
Sex	Embarked	
female	C	64
	Q	27
	S	140
male	C	29
	Q	3
	S	77

OBSERVATION: People who boarded the ship from Southampton seems to have higher chance of survival in both Male and Female

EDUCATED GUESS: Southampton could be an affluent place and thus passengers boarding could have gone for the first class