Contoso

MAKING TOKYO CITY'S NEIGHBOURHOODS COMMERCIALLY VIABLE WITH DATA SCIENCE IN PYTHON

MUHAMMAD SHEHERYAR NAVEED

SKYEARTH

Japan as a Business Hub

- Japan is the third largest economy in the world just behind China and US and makes 6% of the total Gross Domestic Product

- Japan is considered a gateway to Asian Market and provides abundant business ventures.

- Japanese government is quite supportive in terms of helping setup a business. Japanese government lowered down the minimum capital for opening business from 10 million to just 1 million yen in 2003.

- Japan has a highly educated workforce thus making the employees highly devoted to their companies and striving towards the success.

- Tokyo has 75% of foreign companies located and provides incentives to the companies such as subsidies along with a preferential tax system.

EASE OF DOING BUSINESS IN JAPAN

Japan is ranked 29 among 190 economies in the ease of doing business, according to the latest World Bank annual ratings. The rank of Japan improved to 29 in 2019 from 39 in 2018.
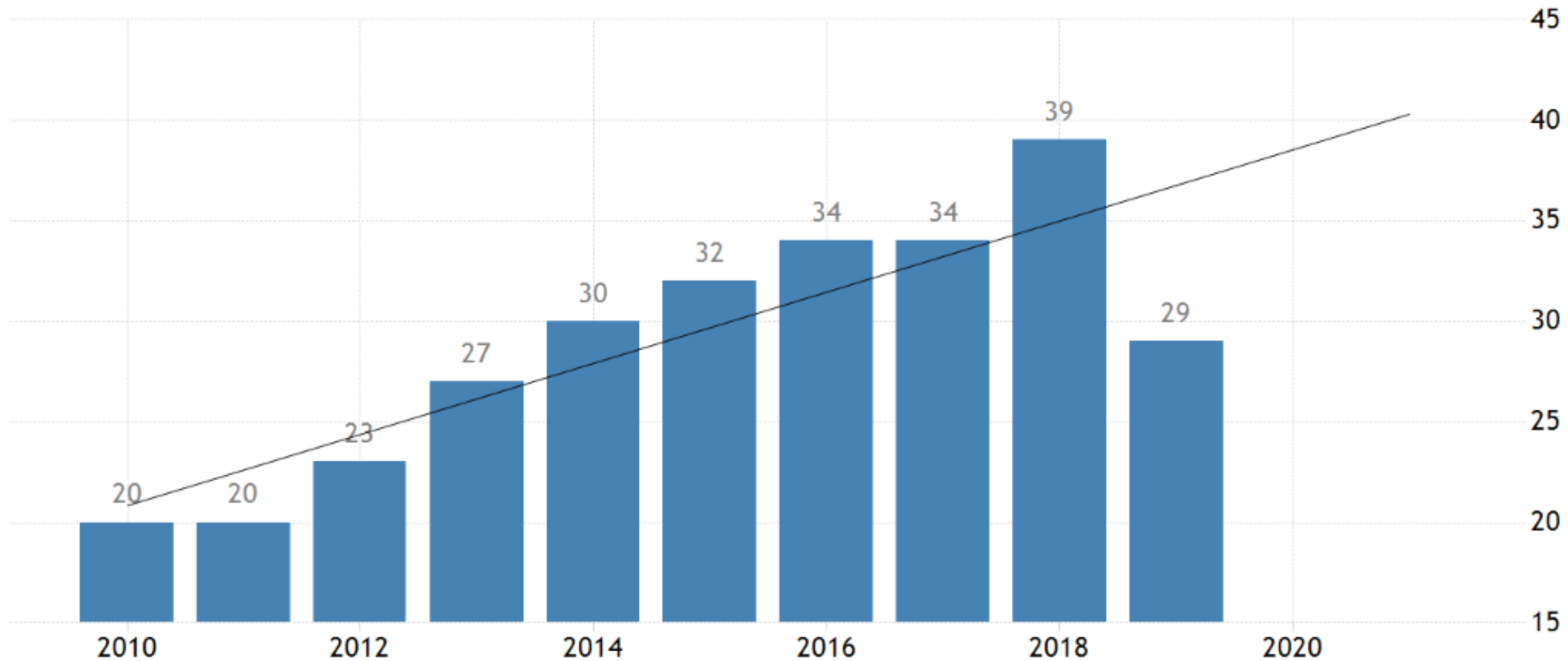
SOURCE: TRADINGECONOMICS.COM | WORLD BANK

Contoso

The Problem

Productivity

Too many factors to consider at any given time and difficult to track the progress.

User Frustration

Things get difficult when the information pool becomes too large and when reliable information is too difficult to obtain.

Scarce Analytics

Difficult to combine/analyze information gathered and make a relation with parameters

Solution

An platform to automate the search for a potential business opportunity based on a given dataset at any particular location.

Unique

Tested

Authentic

Mechanism

Data Science - Python

STEP 1: OBTAINING AND CLEANING DATA

Data Source

- The data was obtained from the following source: *https://sites.google.com/site/yangdingqi/home/foursquare-dataset*

- It consists of data of Tokyo city listing the venues and their respective customer visits accompanied by date time stamps. The data is for ten months from 12 April 2012 to 16 February 2013. It has around 573,703 records for Tokyo.

Refinement

- We need to refine the data obtain just the useful data elements so we dropped the useless columns.

	User_ID	Venue_ID	Category_ID	Category_Name	Latitude	Longitude	Time Zone(off. mins)	UTC Time
0	868	4b7b884ff964a5207d662fe3	4bf58dd8d48988d1d1941735	Ramen / Noodle House	35.715581	139.800317	540	Tue Apr 03 18:22:04 +0000 2012
1	114	4c16fdda96040f477cc473a5	4d954b0ea243a5684a65b473	Convenience Store	35.714542	139.480065	540	Tue Apr 03 19:12:07 +0000 2012
2	868	4c178638c2dfc928651ea869	4bf58dd8d48988d118951735	Food & Drink Shop	35.725592	139.776633	540	Tue Apr 03 19:12:13 +0000 2012
3	1458	4f568309e4b071452e447afe	4f2a210c4b9023bd5841ed28	Housing Development	35.656083	139.734045	540	Tue Apr 03 19:18:23 +0000 2012
4	1541	4b83b207f964a5202c0d31e3	4bf58dd8d48988d1f8941735	Furniture / Home Store	35.705074	139.619502	540	Tue Apr 03 19:20:09 +0000 2012

	Venue_ID	Category_Name	Latitude	Longitude
0	4b7b884ff964a5207d662fe3	Ramen / Noodle House	35.715581	139.800317
1	4c16fdda96040f477cc473a5	Convenience Store	35.714542	139.480065
2	4c178638c2dfc928651ea869	Food & Drink Shop	35.725592	139.776633
3	4f568309e4b071452e447afe	Housing Development	35.656083	139.734045
4	4b83b207f964a5202c0d31e3	Furniture / Home Store	35.705074	139.619502

STEP 2: FINDING THE IN-DEMAND BUSINESS

Process

- We further group all the categories as per the number of visits to obtain the visitor count of each business type.

- There are some categories in the dataset which are out of our interest. These include Train Station. *One cannot consider opening a train station as a business opportunity.*

- Such unwanted categories were removed to obtain a further refined data.

- On sorting the refined dataset in the descending order, we came to know the most in demand business based on the number of customer visits. *In this case, it was 'Electronics Store'*

	Venue_ID	Category_Name	Latitude	Longitude	Visitor_Count
162	4b19f917f964a520abe623e3	Train Station	35.698596	139.773018	160113
322	4b380ad7f964a520bc4a25e3	Subway	35.665115	139.712459	29302
1091	4b55670ff964a52071e327e3	Electronics Store	35.699015	139.774622	6644
2005	4bbac8b753649c742f7249fb	Office	35.699882	139.772414	3167
1872	4b0a57d1f964a5205d2323e3	Mall	35.532942	139.695765	2745

	Venue_ID	Category_Name	Latitude	Longitude	Visitor_Count
1091	4b55670ff964a52071e327e3	Electronics Store	35.699015	139.774622	6644
1990	4b0bc75ff964a520923323e3	Bookstore	35.700427	139.771752	2503
51	4b5d2ad2f964a520535529e3	Convenience Store	35.552431	139.647492	2137
2362	4df8690581304987d7fadc27	Coffee Shop	35.617692	139.728692	1561
995	4b5bc2e0f964a520ad1429e3	Food & Drink Shop	35.770128	139.660808	1478

STEP 3: SEARCHING THE POTENTIAL NEIGHBORHOODS

Process

- For each coordinate we find the number of electronics stores present within its 3kms radius. Then we came across the coordinate that has the most number of electronics store within its 3km vicinity

- As per the results on the right we have coordinates with 21 electronics shops so we chose the coordinate with 20 shops and find its neighbour as our potential target.

```
(35.68435257369705, 139.73719954490662) : 21
(35.68433296624852, 139.74218845367432) : 21
(35.685535547509154, 139.77484703063965) : 21
(35.6864766854162, 139.72904562950134) : 21
(35.686439390552145, 139.68487850418327) : 21
(35.68684072825564, 139.7741263290934) : 21
(35.6847442424678, 139.76341348869008) : 21
(35.686276258829416, 139.73653435707092) : 21
(35.685936700282824, 139.6158176667451) : 21
(35.68668800420691, 139.72854807972908) : 21
(35.68607583173909, 139.77600574493408) : 21
(35.68485310665427, 139.7838270664215) : 21
(35.686038796243366, 139.66934233903885) : 21
(35.686476, 139.784404) : 21
(35.68654857743877, 139.78344082832336) : 21
(35.68440489706805, 139.68536204544893) : 21
(35.686039050003295, 139.7390273) : 21
(35.68681349, 139.773637) : 21
(35.68560219869659, 139.7835977691558) : 21
(35.67887319366422, 139.78747487068176) : 20
(35.679467991491705, 139.74145084619522) : 20
(35.678823082289455, 139.66568380594256) : 20
(35.67837643603711, 139.78663802146912) : 20
(35.68835114950396, 139.76315891999548) : 20
(35.688207767271784, 139.69623030487722) : 20
(35.687082319216536, 139.76545929908752) : 20
(35.688001653824074, 139.69898343086246) : 20
(35.6781683633928, 139.7681871056557) : 20
(35.683384000000004, 139.70154385) : 20
(35.68161184137628, 139.78618204593658) : 20
Coordinate that has the given specific shop the most:  (35.68435257369705, 139.73719954490662)  which is:  21
```

FINAL RESULTS

ANALYSIS

- The red dot here in the diagram specifies the target location which is attracting the most number of customers in this area.

- The blue markers are the following locations in Tokyo:

  1. Shimbashi
  2. Roppongi
  3. Shinjuku
  4. Ginza
  5. Shinjuku Ward

- These represents the potential locations where a similar business could be initiated to get a good market by facilitating the potential customers. These locations are within 3km of the target location.
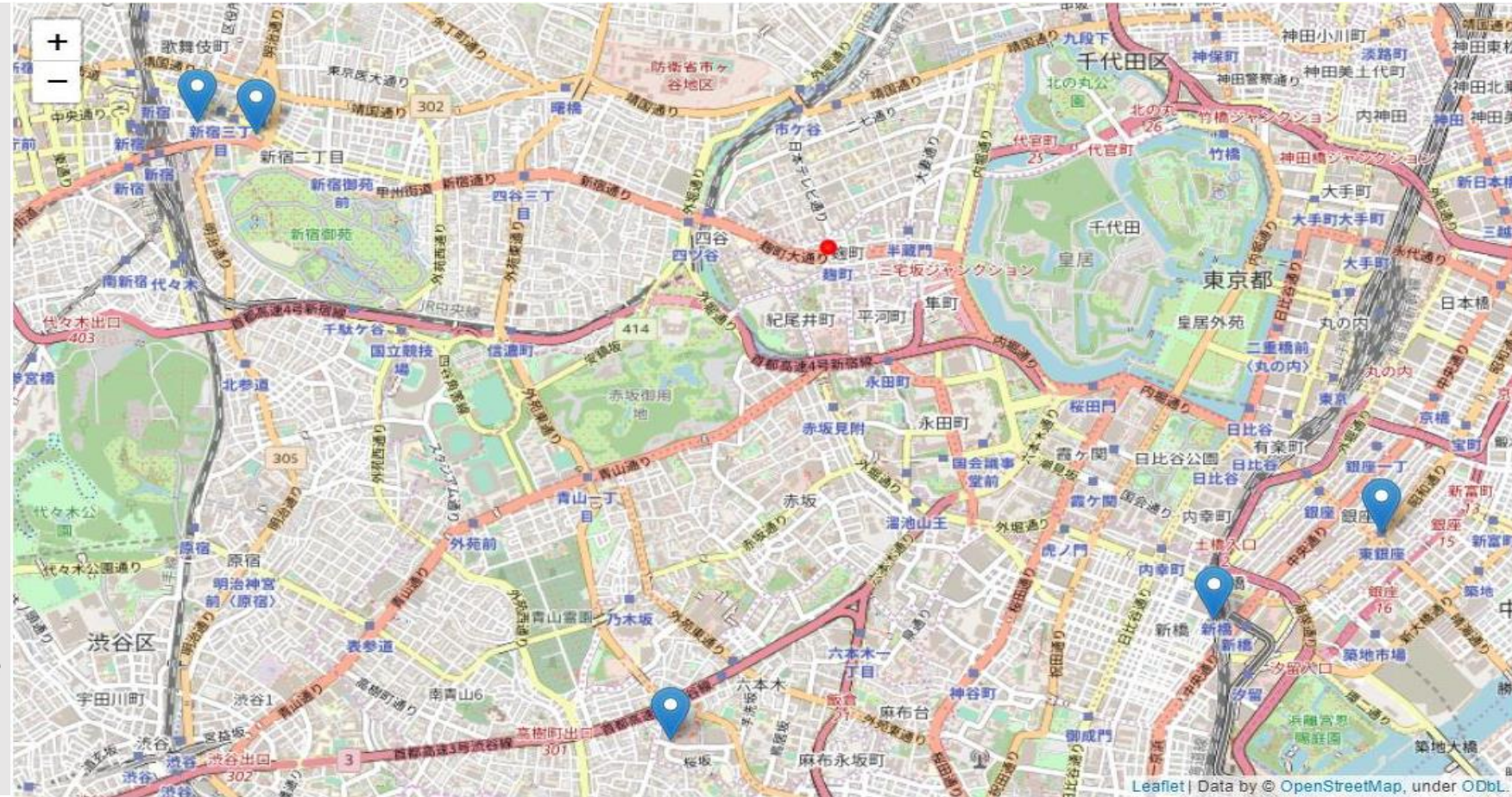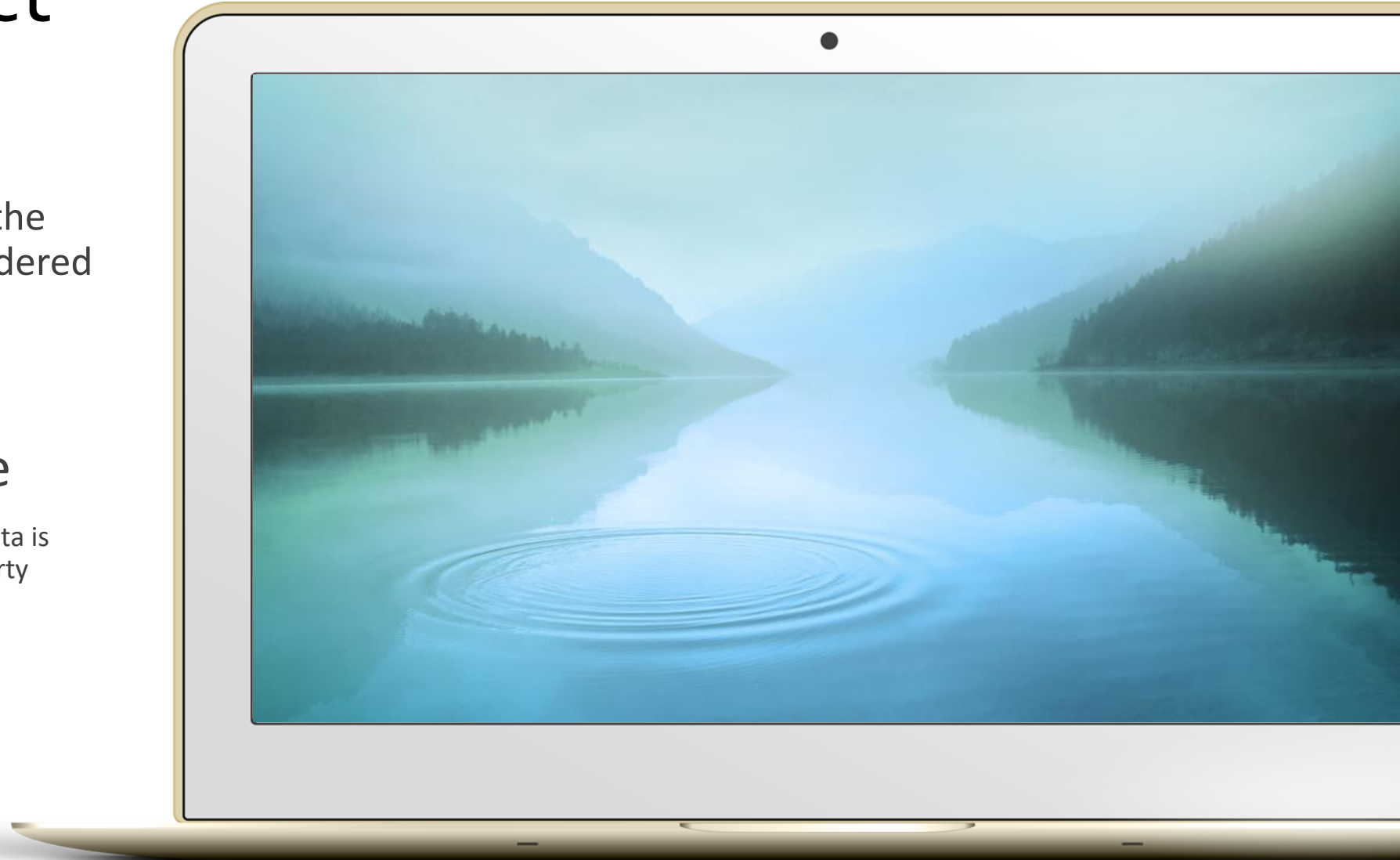
Digital Product

All the hassle is over when all the metrics that needs to be considered are confined on a single digital platform

Other benefits include

- Accurate information as the source of data is known and no more reliance on third party

- Less time consuming.

Thank You!