

INTERNSHIP REPORT

Academic Year: 2020



DOJO
SPORTS

SUBMITTED BY: SUPRIYA RANI

TENURE: 27th February 2020 to 28th August 2020

DURATION: 6 months

DESIGNATION:

Data Analyst Intern (27th February 2020 to 27th May 2020)

Data Analyst Team Leader (28th May 2020 to 28th August 2020)

ORGANIZATION: DOJO Sports Private Limited

Designation 1: Data Analyst Intern

DURATION: 3 months

TENURE: 27th February 2020 to 27th May 2020

SKILLS: Web Scrapping, Python, Microsoft Excel

WEB TECHNOLOGIES: Google Sheets, Jupyter Notebook

Task Involved:


- Data Collection from Websites:

Different websites like Google Maps, Justdial contains ample of data having location related to various sports classes/clubs prevailing all over India/world-wide.

Depending upon the assigned task and its requirement, I had collected data over various attributes as mentioned over the Google sheet using different technologies which includes Python, Web scrapping Techniques.

After collecting the required number of data from the mentioned websites, I used to copy those datasets to the Google sheet whose link is already mentioned through the email.

- Data Collection from Facebook



Facebook is one of the most commonly used social platform worldwide. It is an evolved community now-a-day. This properly suggests that Facebook contains huge source of data.

My one of the assignments also constitute of collection of data from Facebook. Through different Facebook sport groups/ pages, I used to collect their websites, phone number or other contact details to serve our purpose.

- Attributes collected through Website includes:
 - Club/Academy/Class Name
 - Address
 - Google Link
 - State
 - Contact Number
 - Other Contact Number
 - Head/Owner Contact Number
 - Head/Owner Name
 - Membership Plans
 - Batches Timing
 - Instructor Name
 - Instructor Photo
 - Coach Details
 - Notes

- Attributes collected through Facebook includes:
 - Facebook Group Name
 - Facebook Group Link
 - Number of members in the group
 - Last activity status of the group
 - Website URL

- Contact Number
- Group location
- Any other details

Designation 2: Data Analyst Team Leader

DURATION: 3 months

TENURE: 28th February 2020 to 28th August 2020

SKILLS: Python, Microsoft Excel, Data Management, Data Analysis, Predictive Analysis

WEB TECHNOLOGIES: Google Sheets, Jupyter Notebook, Google Meet, Microsoft Word, Microsoft Presentation, Microsoft Outlook

My data collection dataset includes different sports like Badminton, Volleyball etc. and other data about Facebook groups and its pages.

Task Involved:

- **Candidate selection for the team:**

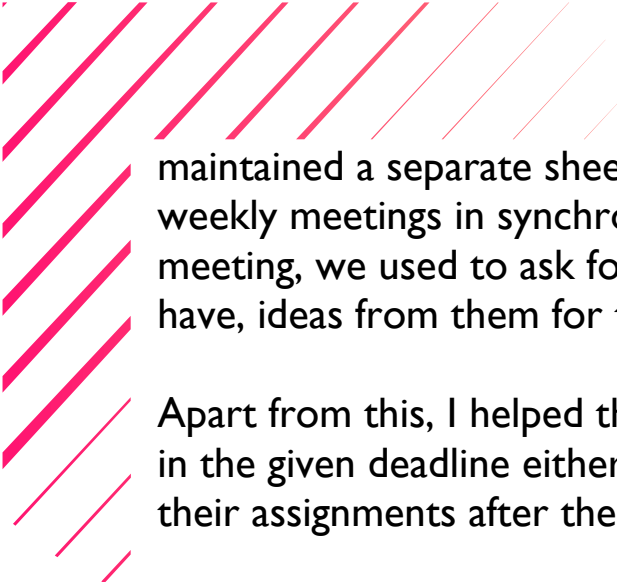
Everything starts with the right candidate in the team. Candidates are the bricks of the building. I have selected 6 interns in my team through telephonic interview, resume selection and update the same in the database.

Link:

https://docs.google.com/spreadsheets/d/1uxtbzl8ezCi0o9c_V0aEGuvq1ZQ0xKf1Jopxz3h68U8/edit

- **Task Management for Interns:**

Firstly, I have synchronized the task to be assigned to interns. “DOJO Timeline” sheet was used for this purpose. I have also



maintained a separate sheet for every intern. I also used to call weekly meetings in synchronization with ML department. In this meeting, we used to ask for their work progress, queries if they have, ideas from them for the betterment of the company.

Apart from this, I helped them in their task, if they unable to finish in the given deadline either by ideas, technologies or completing their assignments after the deadline.

Links of the intern's sheet are as follows:

DOJO Timeline:

<https://docs.google.com/spreadsheets/d/1DMQi7g2MQOBVWdOj6gtGKyRE3-OgcKCARIePHq6x848/edit#gid=1221245668>

DOJO Interns sheets:

Sakshi Karn

<https://docs.google.com/spreadsheets/d/16iYVUPROLilcZ3yylPpvaXQQ7FrBzzvD18R4FxFhcOU/edit#gid=0>

Ekta Sharma

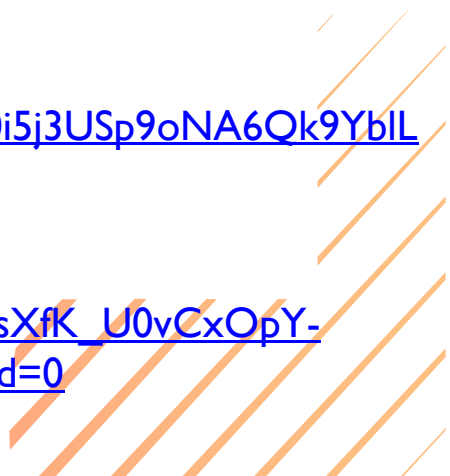
<https://docs.google.com/spreadsheets/d/17alwzo6b52ClrwggLRZWnNOkHNVrFmhRHiA48O2GjNg/edit#gid=0>

Chenchala Sai Pravallika

<https://docs.google.com/spreadsheets/d/1H0i5j3USp9oNA6Qk9YbLsK-E74q0sE27Rl642ZFRGw/edit#gid=0>

Abhishek Khanna

https://docs.google.com/spreadsheets/d/123sXfK_U0vCxOpY-9R4XIMPVft7BetNV6oQNIHKLsEk/edit#gid=0



Sarah Stalin

https://docs.google.com/spreadsheets/d/1wFg4bVQpRVNnZmDWJA VpK7mz5AJ0uou5ssC_IILqf9U/edit#gid=0

Above interns has started and completed their internship in my tenure totally, while the interns listed below are managed by me up to more than half of their internship duration.

Aaditya Rana

https://docs.google.com/spreadsheets/d/1JAp_jlPZPL4ufnGApJw2uoB SQIVOdYqwmIUfYi8YuG0/edit#gid=1969936879

Bhavika Aggarwal

https://docs.google.com/spreadsheets/d/1CskkjWMneaMtR_WhCVt5 9yA8SBH4ANadpEfQ2GkONGs/edit#gid=0

Cindy Baretto

https://docs.google.com/spreadsheets/d/1-g8lnjbdak39r2D34pP4s2w66cXZXjdSsrpz65_iFfc/edit#gid=0

Apart from that I used to add the data from these sheets (attribute wise) on daily basis to another sheet to store it in collective manner. It contains the sheets of different sport.

Link:

<https://docs.google.com/spreadsheets/d/1GMr41MAEswJgAbNx-9sDbd70EE3f5EGvuWHtHWRNvqQ/edit?ts=5ef5b662#gid=430671174>

Also, I used to manage the daily progress sheet of the interns on a sheet.

Link:

https://docs.google.com/spreadsheets/d/13zQB5AdmOO0_diMmIJG BI8vjwGhcepp-J8MXDWGEsY/edit#gid=0

DOJO ANONYMOUS

Contains sheets of Akshita Aggarwal(lost the interest), Shourya Pratap Bhadauriya(completion of internship duration), Abhayjyot Singh Grover(left in the midway), Rishabh Ralli (inducted for Machine Learning department, data collection is assigned as the beginning task)

Some of the data collection has occurred in this page. Reason behind this is some of the interns used to left in the meanwhile because of their personal reasons like laptop issue, unavailability, lost the interest, completion of internship tenure. So, their dataset is properly managed in one sheet and updated the same in the main sheet.

Link:

<https://docs.google.com/spreadsheets/d/12AdBTPRCWInudVXgCtKrLztFhmLr2K-cJl7d5OhRTz4/edit#gid=0>

- **Task completed by me (apart from the management):**

- I have scrapped the data related to turfs from different websites, some of which are suggested by my superiors and others are surfed by me. After scrapping the data, there was irregularities, so I properly managed them with the help of my
- interns and update the same in the below sheet named as “Turf Organized data”.

Link:

https://docs.google.com/spreadsheets/d/1FE5ndVWX3Hlbq_oHPEZaEI6tHQaY8EGoO3eYTuaOHl-s/edit?ts=5f0bfada#gid=694484851

○ OPENING OF CLUBS PREDICTIVE ANALYSIS

CORONA is prevailing everywhere and this has changed the scenario of market. Even sports dependent industries too are badly affected to it. Through the above analysis, I had predicted the most probable states that will allow the reopening of clubs. This prediction was totally dependent on CORONA cases, norms of the government, population and population density etc.

I had submitted this report to the superiors in the form of word file.

○ OLYMPIC DATASET ANALYSIS

Olympic datasets are analyzed specifically for India, to get the answer of some important questions that includes:

✓ Which sports has bragged largest medals?

✓ In which sport, highest number of athletes are participating?

Above questions can help us to build the strategy to focus.

- Analysis are performed time to time to answer the questions of superiors as well as for my personal development of strategy.

Some of the major analysis subjects are:

- ✓ State wise famous cities of India and how many of them are already present in our sheet?
- ✓ Excel sheet having all the sports but the clubs/classes/academy should only be from South Region.

- ✓ Number of clubs we have till so far city/state wise.
- ✓ Datasheet containing the datasets having no contact details either email/ contact number/ fax number.
- ✓ Sport wise analysis of total number of data till date.

○ DOJO PLAYBOOK FOR DATA ANALYSIS TEAM

I have designed a DOJO Playbook for Data Analysis Team for efficient workflow inside the team based on the ideology of superiors. This playbook/ manual can brief the working process of the team with its objectives, so that a new person whenever starts working for the team can get the brief idea. This book also contains the name of individuals responsible for the particular task which is changed time to time.

○ MARKET ANALYSIS

Some of the competitors' website are also analyzed by me to have any clue from them for our development. I have also used this to understand their mindset and working strategies behind the organization.

○ ANALYSIS OVER E-MAIL FILTERING AND GOOGLE WEBSITE POSITIONING / RANKING SYSTEM

E-Mail filtering techniques are studied and analyzed which is useful for Marketing team, so that they can push the email in such a manner that it won't get filtered out.

In google positioning system, I have tried to find out which factors are responsible for giving higher position to any website in the google page and suggested the same to the team. I have documented both the analysis results into one PDF file.



○ HEVO WEBSITE ANALYSIS

I have analyzed website “HEVO”. The purpose behind this to get the information if they can help in our needs that fits us financially.

○ TURFS DATA ANALYSIS

I have performed the analysis over the turfs data that are collected by me to get an insight about it. The major questions that were answered through this analysis includes:

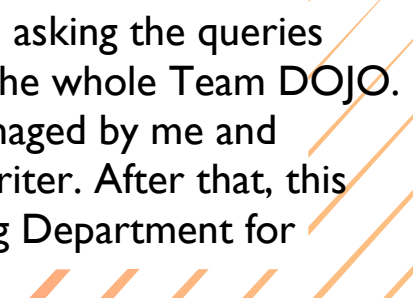
- ✓ Which states are present in our data?
- ✓ Which state has maximum number of turfs?
- ✓ How many sports are there in the data?
- ✓ Which sport has maximum turfs all over India?

○ CLUBS DATA ANALYSIS

The data that are collected by the fellow team members are properly synchronized with the given sheet's attributes. Apart from them I used to perform analysis over that to answer the questions of my superiors. I have listed some of the questions listed above.

○ CHATBOT DATA

One of intern has made a google form asking the queries consisting of 3 major questions from the whole Team DOJO. These collected data are properly managed by me and completed with the help of content writer. After that, this dataset is handled to Machine Learning Department for generating Chat-Bot.



LINK:

https://docs.google.com/spreadsheets/d/1-PBIO_9e9CqZ56wjxT6lvPQ-NKF4GaO59CBkrkvy6QQ/edit#gid=0

NOTE: All the above links are accessible to dojosportsofficial@gmail.com . In certain cases, other team members are accessible to it.

SUMMARY (WORK EXPERIENCE)

I really enjoyed the work culture of DOJO Sports Private Limited. Shivali Thakur (HR) and Akshay Bavkar(CEO), both are very supporting and great person. Sugam Thale (COO) is a bit introvert person but a lot brainy by nature. His work speaks louder than his words. My fellow team mates Shikhar Rajput is very enthusiastic and charming personality.

Other interns of this family are quite amazing and talented. I wish I would have this type of talented and supportive minds in the future as well as the work culture.

I can't believe that I have completed 6 months here. Especially the idea and implementation of fun session is quite amazing.

Apart from this, the practical knowledge that I have acquired from here will be helpful for me in the future. I am really overwhelmed by happiness that you chose me as your team leader.

I hope I had met your expectations.

Thankyou!!
Supriya Rani