Python Analysis of Linkedin Social Saturday

1.0. Background

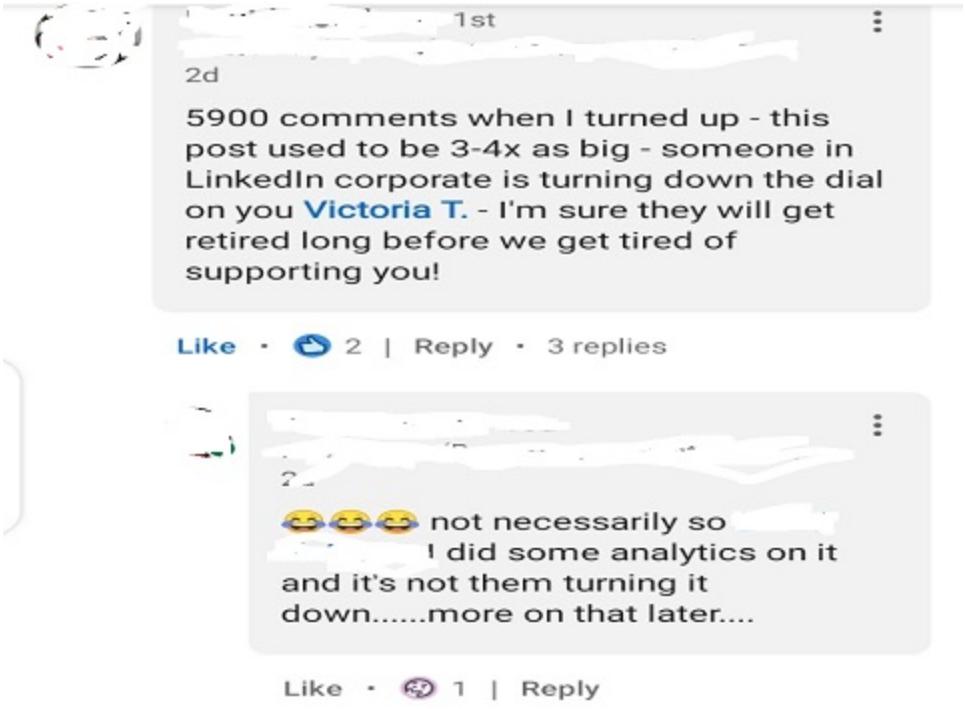
Social Saturday is an open LinkedIn community/party where LinkedIn users gather to socialize and make new connections, with the intention of nurturing those connections to mutually beneficial (corporate) relationships.

The community originated from a LinkedIn user with the handle 'Jonathan Lee', but it was popularized by another Linkedin user - Victoria T. The community thrives on the power of network externalities - the more people join, the more valuable the community becomes. There are no requirements for starting a Social Saturday Party and any member can start one by using the the hashtag #socialsaturday. Since Victoria's community was the most visited, this analysis focuses on her on her community.

Every Saturday, LinkedIn users who intend to expand their network with minimum effort gather at Victoria's page and interact with her Social Saturday party post. Users introduce themselves and invite other users to connect with them. My interactions with the parties contributed to a 495% growth in my network within 4 weeks, as it grew from 202 connections to 1000.

However, after 4 weeks of participation, diminishing returns were observed. Less people were showing up and interacting with Victoria's party, while a good number of the members were familiar connections. Members expressed complaints that the platform had a hand to play in the decline the .

Fig 1: Member Complaints about Decline



Towards this end, data mining and exploration was done to understand the historical performance of the community and know what the data had to say about it

This report therefore provides:

- The observations from the analysis.
- A possible explanation for the decline in performance.
- Recommendations to improve the community.

Most importantly, this report could also serve as a platform-agnostic template for community builders who intend to build long-lasting communities.

1.1. Web scraping and data extraction.

That data was extracted from two sources:
1.Victoria's social saturday posts.
2.Search results using the hashtag '#socialsaturday'

- The data was extracted on the 5th of October, 2022 and covered a period of 25 Saturdays across 6 months, from the 3rd Saturday in April, 2022 to the first Saturday in October, 2022.
- A Linkedin search using #socialsaturday produced 300 parties which consisted of groups, individual posts and polls. However, the search result did not deliver a complete list.
- The numbers of Likes, shares and comments on Victoria's social Saturday posts were captured and recorded in CSV files.

1.2. Libraries used

Pandas library: Pandas was used to convert the datatype from CSV to DataFrame, using the pd.read_csv function, and loaded into the Jupyter notebook for further analysis.

Matplotlib: Matplotlib was used to visualize the engagement metrics with bar charts and line charts.

2.0 Conversion to Pandas DataFrame

In [1]:
```
#importing pandas library
import pandas as pd

#converting to Dataframe and removing the index label
df = pd.read_csv('social_saturday.csv')
```

In [2]:
```
df.head()
```

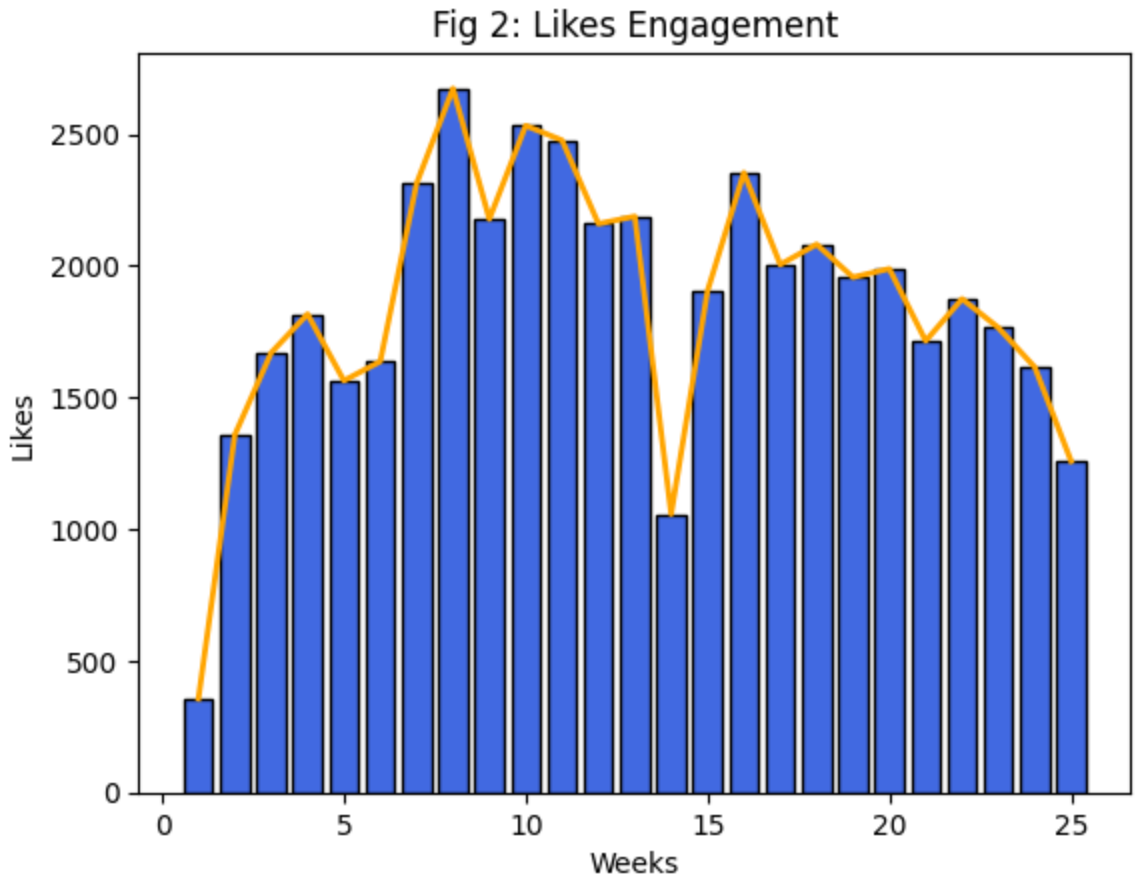
Out[2]:

	Week	Likes	Comments	Shares	Other_parties
0	1	356	535	8	0
1	2	1358	4075	38	0
2	3	1670	6915	101	5
3	4	1817	9531	88	5
4	5	1564	9424	66	10

3.0. Visualization

In [3]:
```
#importing matplotlib, pandas and numpy
from matplotlib import pyplot as plt
import pandas as pd
import numpy as np
```

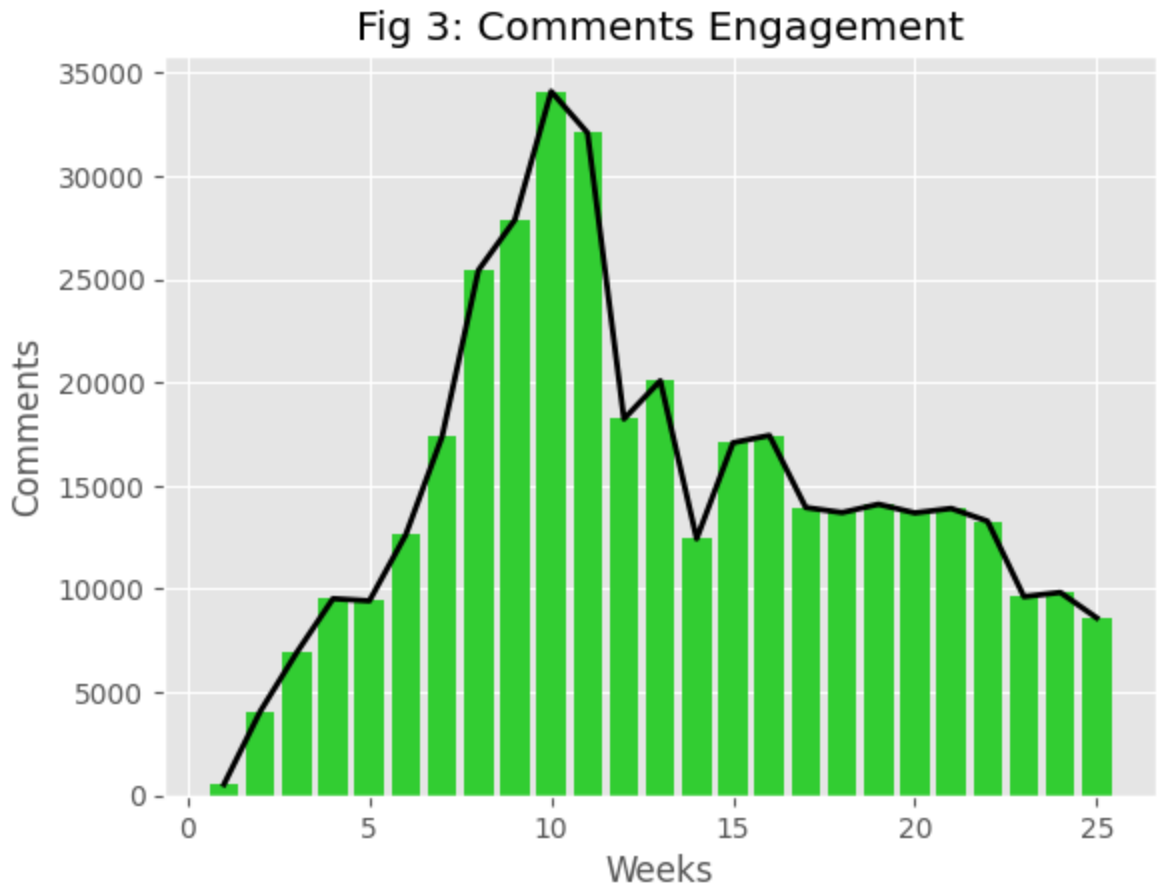
In [4]:
```
# A bar and line chart showing the Likes per week
plt.bar(df.Week,df.Likes, color='royalblue', edgecolor='black')
plt.plot(df.Week,df.Likes, color='orange', linewidth=2)
plt.xlabel('Weeks')
plt.ylabel('Likes')
plt.title('Fig 2: Likes Engagement')
plt.style.use('ggplot')
plt.show()
```



In [5]:
```
# Shows the minimum Likes
print(f"{'Minimum Likes ='}",
      df.Likes.min())

# Shows the maximum Likes
print(f"{'Maximum Likes at peak level ='}",
      df.Likes.max())

# Shows the Likes as at time of data capture
print(f"{'current Likes ='}",
      df.Likes[24])

#Shows the percentage growth at the peak level
print(f"{'Percentage Growth at peak level ='}",
      round(df.Likes.max()/df.Likes.min()*100),'%')

#Shows the percentage drop from peak level
print(f"{'Percentage drop from peak level ='}",
      round(((df.Likes.max()- df.Likes[24])/df.Likes.max())*100),'%')
```

```
Minimum Likes = 356
Maximum Likes at peak level = 2674
current Likes = 1258
Percentage Growth at peak level = 751 %
Percentage drop from peak level = 53 %
```

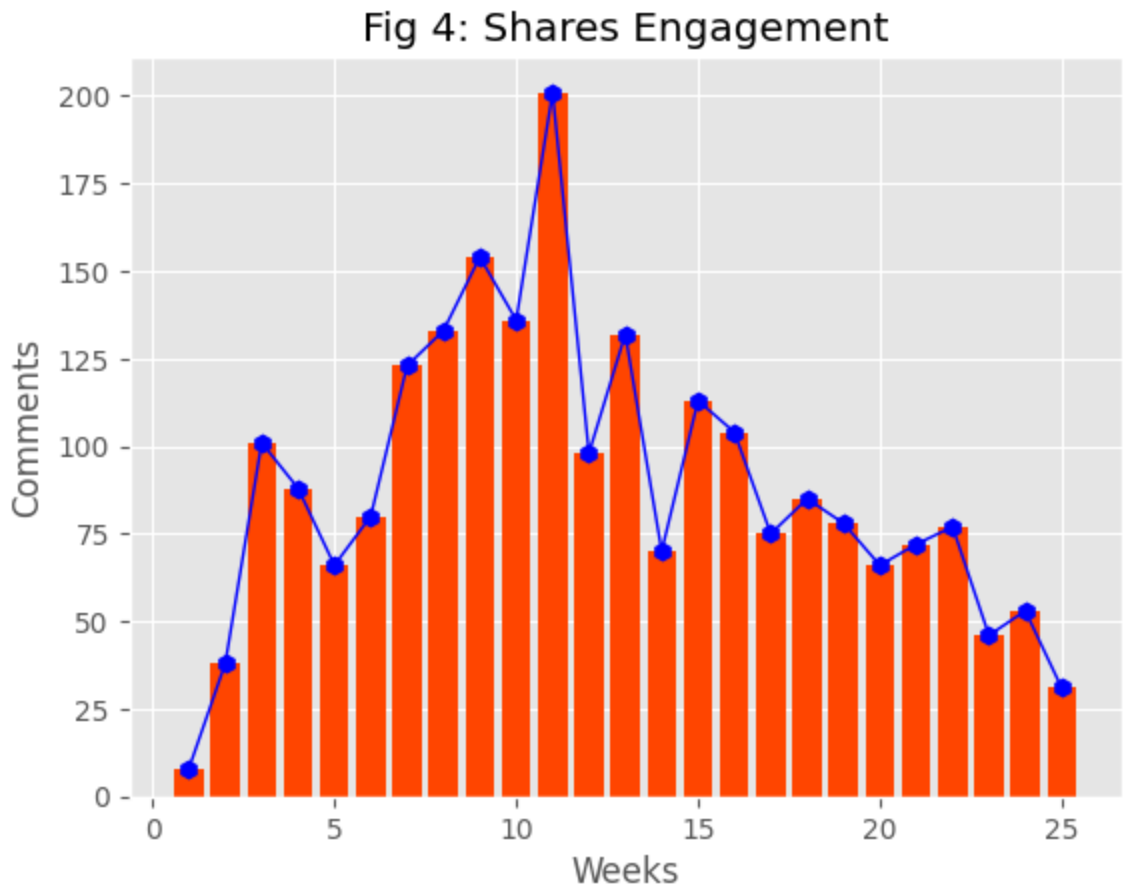
In [6]:
```
# A bar and line chart showing the comments per week
plt.bar(df.Week,df.Comments, color='limegreen')
plt.plot(df.Week,df.Comments, color='black',linewidth=2)
plt.xlabel('Weeks')
plt.ylabel('Comments')
plt.title('Fig 3: Comments Engagement')
plt.style.use('ggplot')
plt.show()
```



In [7]:
```
# Shows the minimum Comments
print(f"{'Minimum Comments ='}",
      df.Comments.min())

# Shows the maximum Comments
print(f"{'Maximum Comments at peak level ='}",
      df.Comments.max())

# Shows the Comments as at time of data capture
print(f"{'current Comments ='}",
      df.Comments[24])

#Shows the percentage growth at the peak level
print(f"{'Percentage Growth at peak level ='}",
      round(df.Comments.max()/df.Comments.min()*100),'%')

#Shows the percentage drop from peak level
print(f"{'Percentage drop from peak level ='}",
      round(((df.Comments.max()- df.Comments[24])/df.Comments.max())*100),'%')
```

```
Minimum Comments = 535
Maximum Comments at peak level = 34097
current Comments = 8601
Percentage Growth at peak level = 6373 %
Percentage drop from peak level = 75 %
```

In [8]:
```
# A bar and line chart showing the shares per week
plt.bar(df.Week,df.Shares, color='orangered')
plt.plot(df.Week,df.Shares, color='blue', marker='h', linewidth=1)
plt.xlabel('Weeks')
plt.ylabel('Comments')
plt.title('Fig 4: Shares Engagement')
plt.style.use('fivethirtyeight')
plt.show()
```

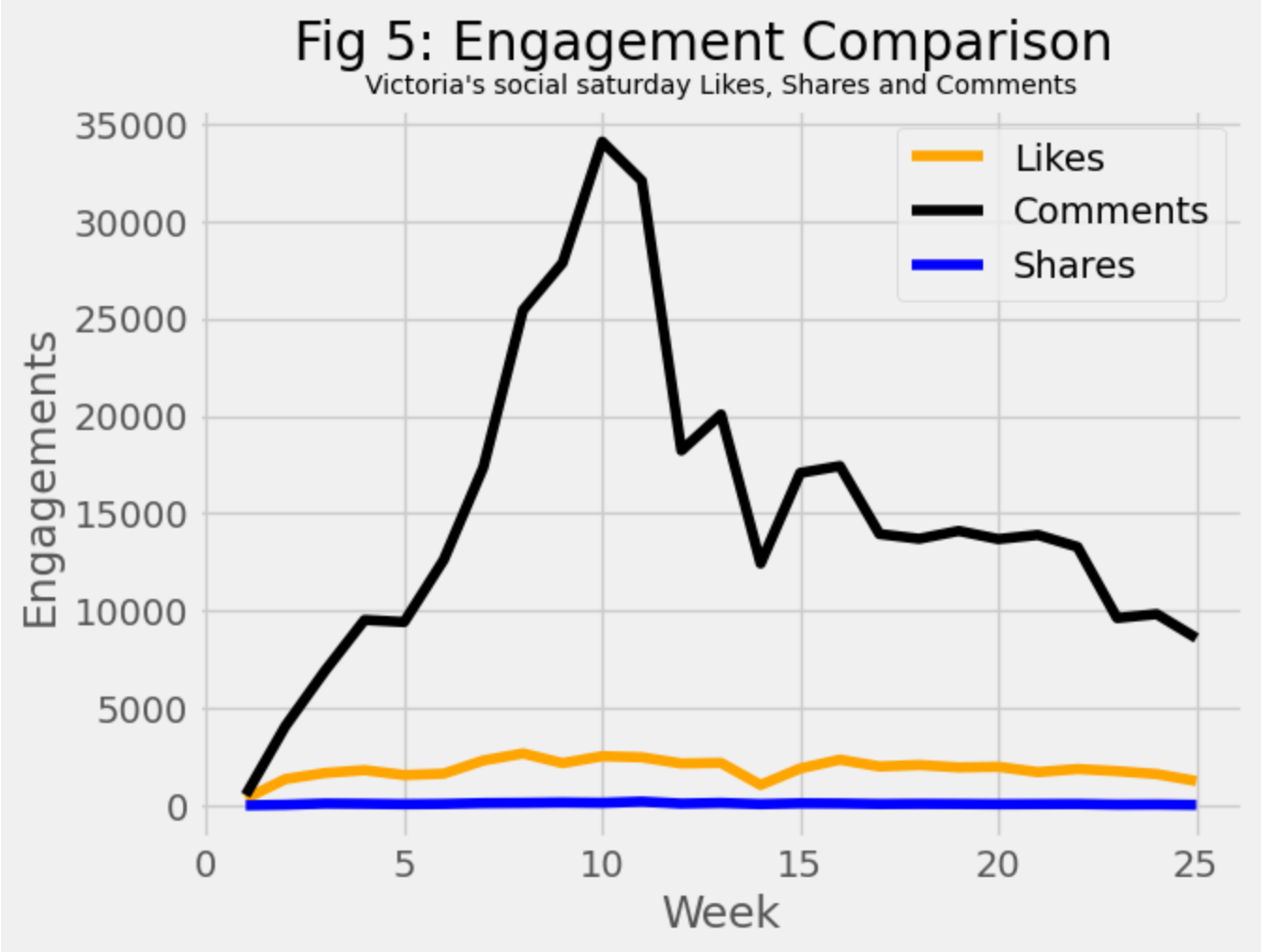


In [9]:
```
# Shows the minimum Shares
print(f"{'Minimum Shares ='}",
      df.Shares.min())

# Shows the maximum Shares
print(f"{'Maximum Shares at peak level ='}",
      df.Shares.max())

# Shows the Shares as at time of data capture
print(f"{'current Shares ='}",
      df.Shares[24])

#Shows the percentage growth at the peak level
print(f"{'Percentage Growth at peak level ='}",
      round(df.Shares.max()/df.Shares.min()*100),'%')

#Shows the percentage drop from peak level
```

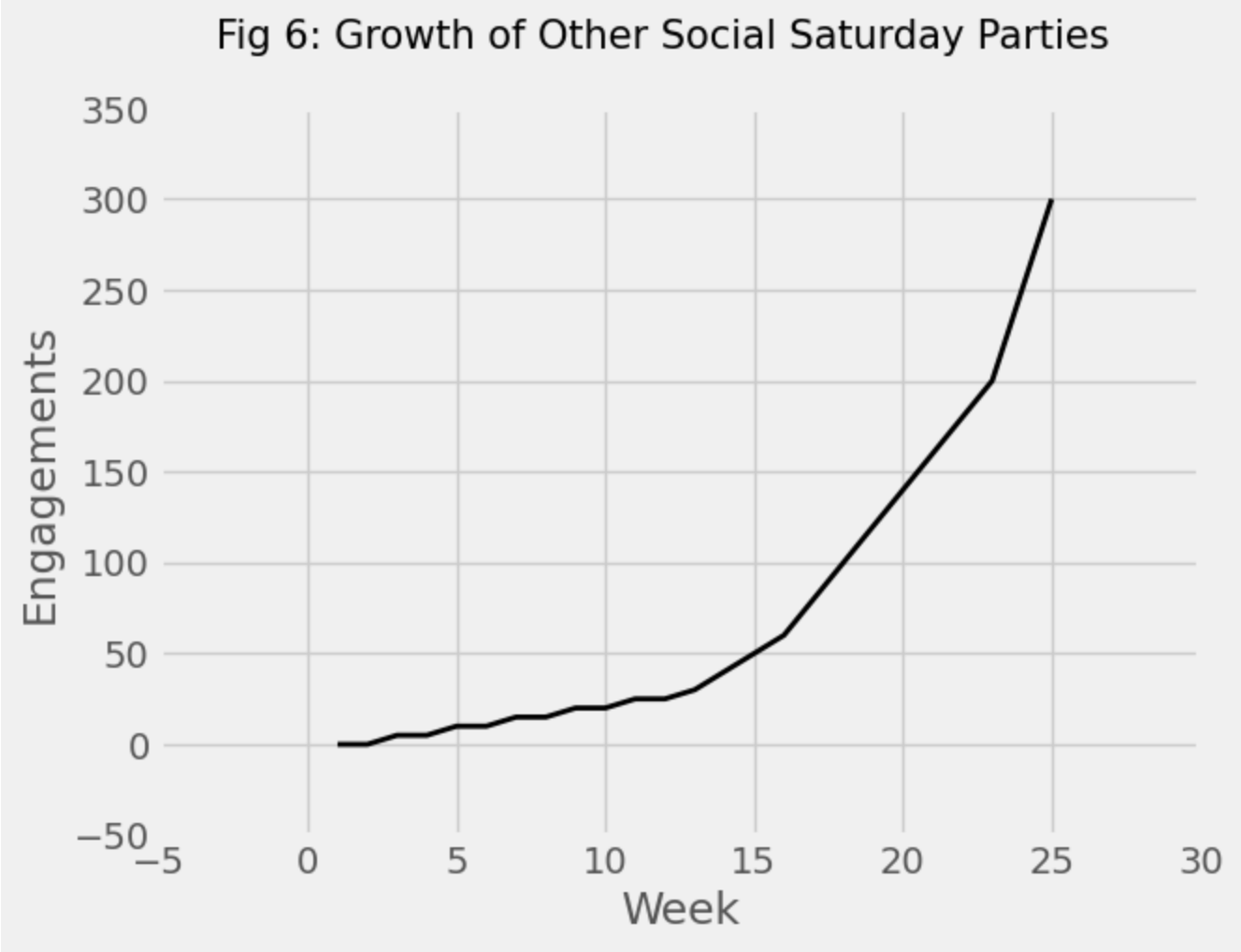
```
print(f"{'Percentage drop from peak level ='}",
      round(((df.Shares.max()- df.Shares[24])/df.Shares.max())*100),'%')
```

```
Minimum Shares = 8
Maximum Shares at peak level = 201
current Shares = 31
Percentage Growth at peak level = 2512 %
Percentage drop from peak level = 85 %
```

In [10]:
```
plt.plot(df.Week,df.Likes, color='orange',label='Likes')
plt.plot(df.Week,df.Comments, color='black',label='Comments')
plt.plot(df.Week,df.Shares, color='blue',label='Shares')
plt.xlabel('Week')
plt.ylabel('Engagements')
plt.suptitle('Fig 5: Engagement Comparison',fontsize=20)
plt.title("Victoria's social saturday Likes, Shares and Comments",fontsize=10)

plt.legend()
plt.show()
```



In [11]:
```
plt.plot(df.Week,df.Other_parties, color='black', linewidth=2)
plt.xlabel('Week')
plt.ylabel('Engagements')
plt.suptitle('Fig 6: Growth of Other Social Saturday Parties',fontsize=15)
plt.style.use('classic')
plt.show()
```



In [12]:
```
print(f"{'Percentage Growth of other parties ='}",
      round(df.Other_parties.max()/1*100),'%')
```

```
Percentage Growth of other parties = 30000 %
```

4.0 Explanation

Week 1 had the lowest engagement with 356 Likes, 535 Comments, 8 Shares.
Peak party attendance was during week 10 with 2533 Likes, 34,097 Comments, 136 Shares.
As at the time the data was taken in Week 25, the engagement had dropped from its peak to 1258 Likes, 8,601 Comments, 31 Shares. This represented a 53 % drop from its peak Likes, a 75% drop from its peak comments and a 85% drop from its peak Shares.
Given the 30000% growth in other parties within the same timeframe, the data shows that fewer people were visiting the main party.
Users were posting multiple times and not everyone was 'Liking' the main post.
Shares were the least used form of engagement.

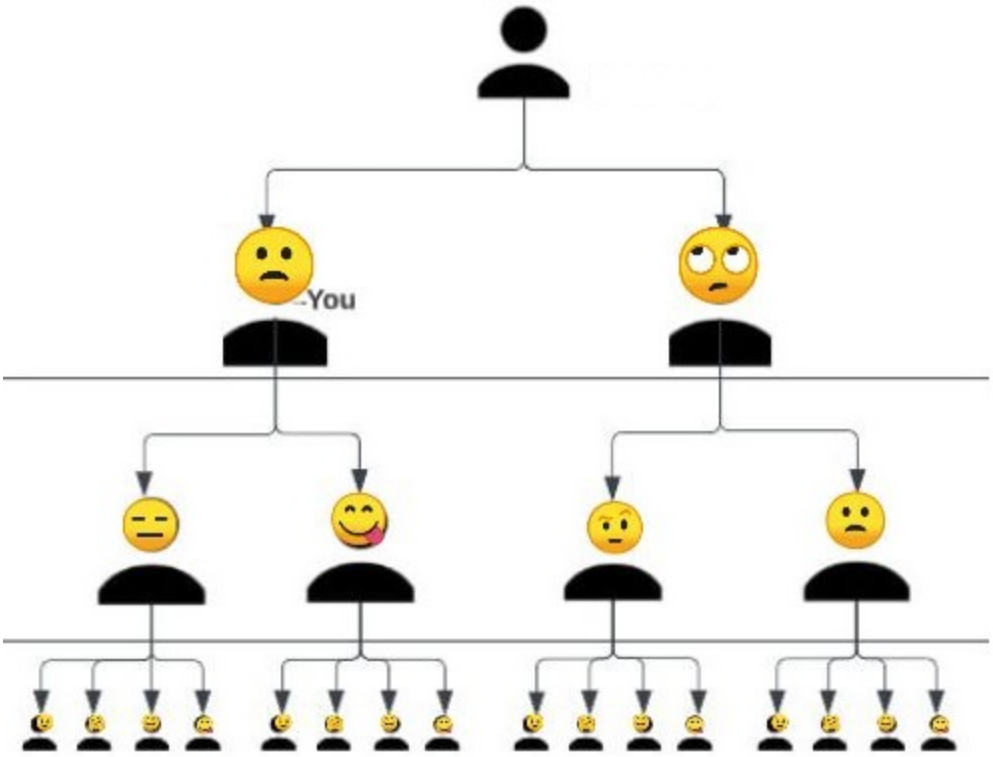
The decline can be explained from two perspectives:

- Growth of other parties and non-tagging of the main party
- The law of diminishing marginal utility
- Product Life Cycle

4.1. Growth of other parties and non-tagging of the main party

The open nature of the parties permits anyone to start a party.
The result shown from the growth of the other parties in Fig x shows that the general party with the hashtag #socialsaturday has experienced tremendous growth within the 6-month period.
Members have branched out to start new parties, but failed to tag the main party to attract people in their parties to the main party.
Directing people from their party to the main party would have created a new pool of second and third connections for people in the main party and extended its life cycle. This attitude continues endlessly and creates a theoretically endless chain of unrelated connections sprouting out of single parent.

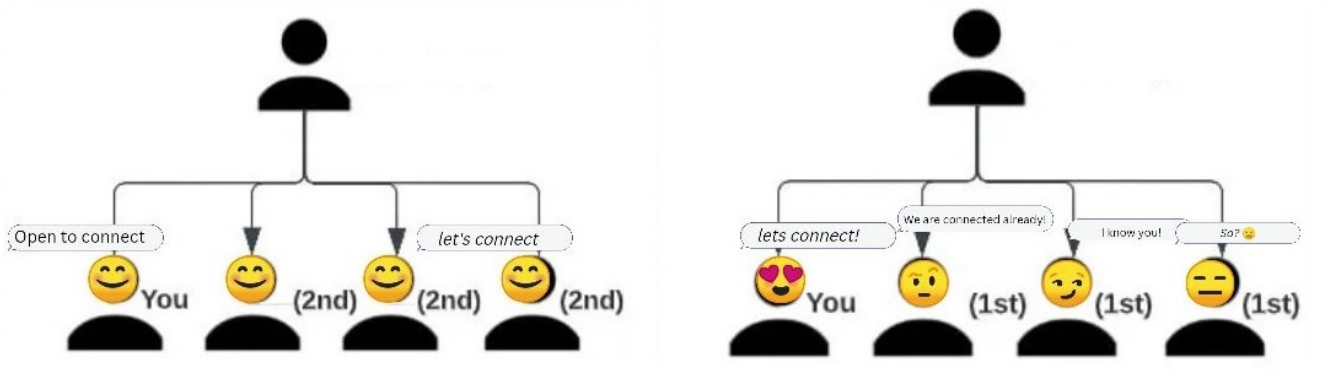
Fig 7: Isolated expansion of family tree



4.2. Diminishing Marginal Utility:

The phenomenon states that the additional utility/value derived from an increase in consumption begins to decrease with each additional input. With respect to the reduction in engagement, this can also be viewed from various angles: The more the same people visit the party, the more familiar everyone becomes and the less valuable the party is with regards to the main objective of meeting new people. All second and third connections eventually become first connections, and there are no new people left. This can be explained with the tree diagram below

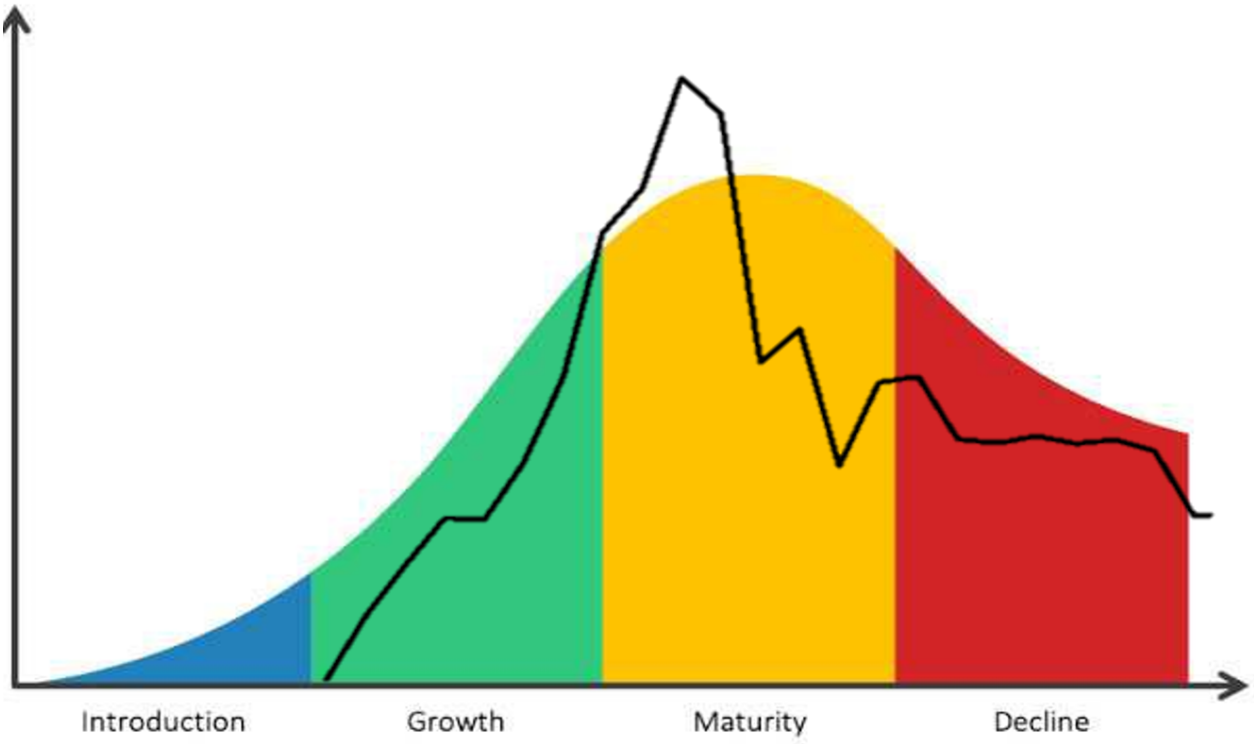
Fig 8: Change in attitude between Week 1 and Week 2



4.3. Product Life cycle

The product life cycle is made up of 4 stages:
- The introduction stage where the community kicks off.
- The growth stage is where the community comes to the knowledge of the public and starts to gain increasing adoption.
- The maturity stage where the community peaks and reaches optimal engagement.
- Decline stage where the party starts to decline.

Fig 9: PLC + Engagement

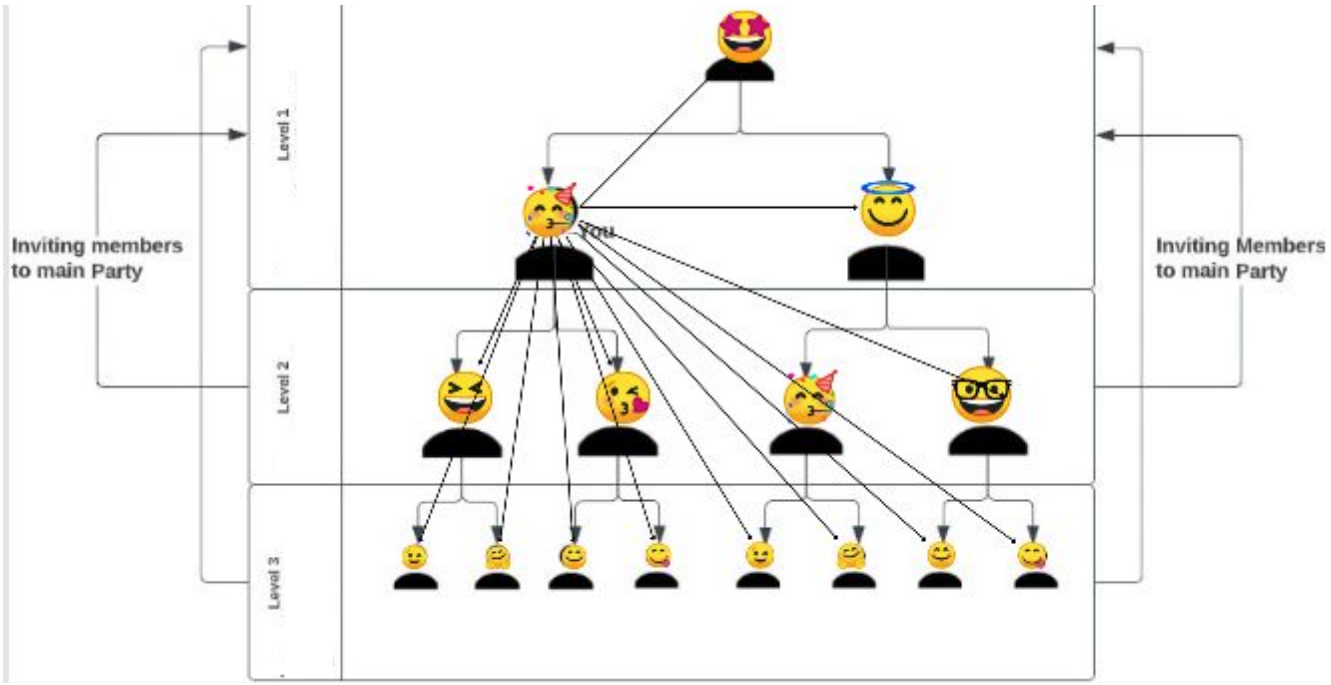


The above image shows how the engagement trend perfectly mimicks the rise and fall of the product life cycle

5.0. Recommendations

- Engaging with posts and party members with likes and personalized comments.
- Tagging the main party or other related parties to leverage the Power of networks
- Creating a themed party. The parties could become monotonous when the theme is simply about making connections and does not promote engagement.
- Inviting other party members to attend the main party and other sub parties by tagging the main parties in order to avoid getting to a party where everyone knows everyone – where everyone is a 1st connection. This would create a constant flow of new members and an endless chain of 2nd and 3rd connections to connect with.
- Following the Social Saturday Squad page and the hashtag '#socialsaturday'.

Fig 10: Interconnectedness



In [ ]: