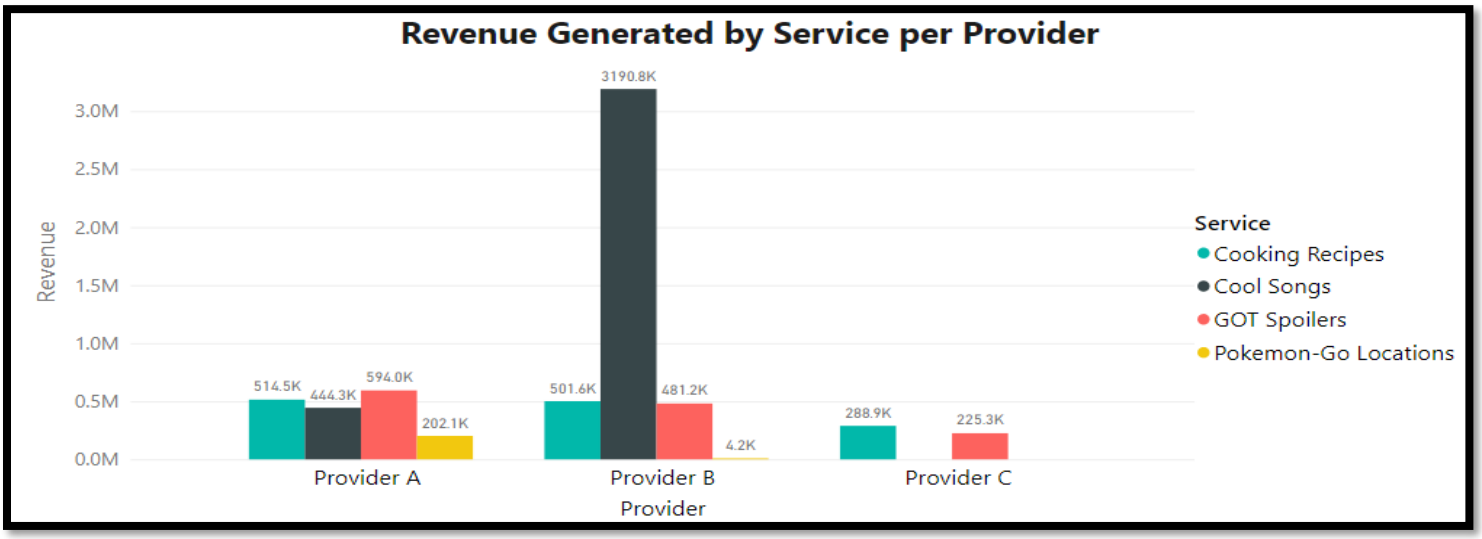
Mncedisi Lindani Mncwabe Submission – Hyve Mobile Data Assessment

Section A

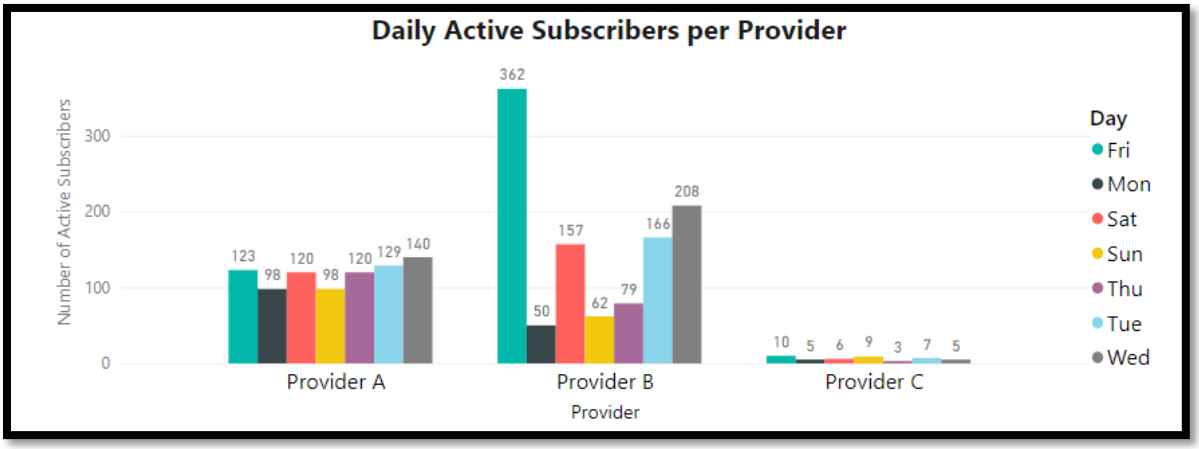
NB: *Performance shown only for successful transactions*

1.)



- **Provider A – GOT Spoilers** account for the most revenue **34% (R594K)** generated by provider A**.** This service is the best performer for this provider.
- **Provider B –** The best performing service for this provider is **Cool Songs**, which accounts for **76% (R3.19M)** of the total revenue generated by this provider.
- **Provider C –** Approximately **56% (R289K)** of the total revenue made by provider C is generated by **Cooking Recipes**, which makes this service the best performer for this provider.

2.)



- **Provider A** sees the most active subscribers on Wednesdays. This day accounts for **17% (140)** of the total active subscribers that provider A gets weekly.

- On the other hand, **Provider B** gets the most active subscribers on a Friday with **33% (362)** of the weekly total active subscribers active on this day.

- **Provider C** also has the most active subscribers **(22% or 10)** on a Friday.

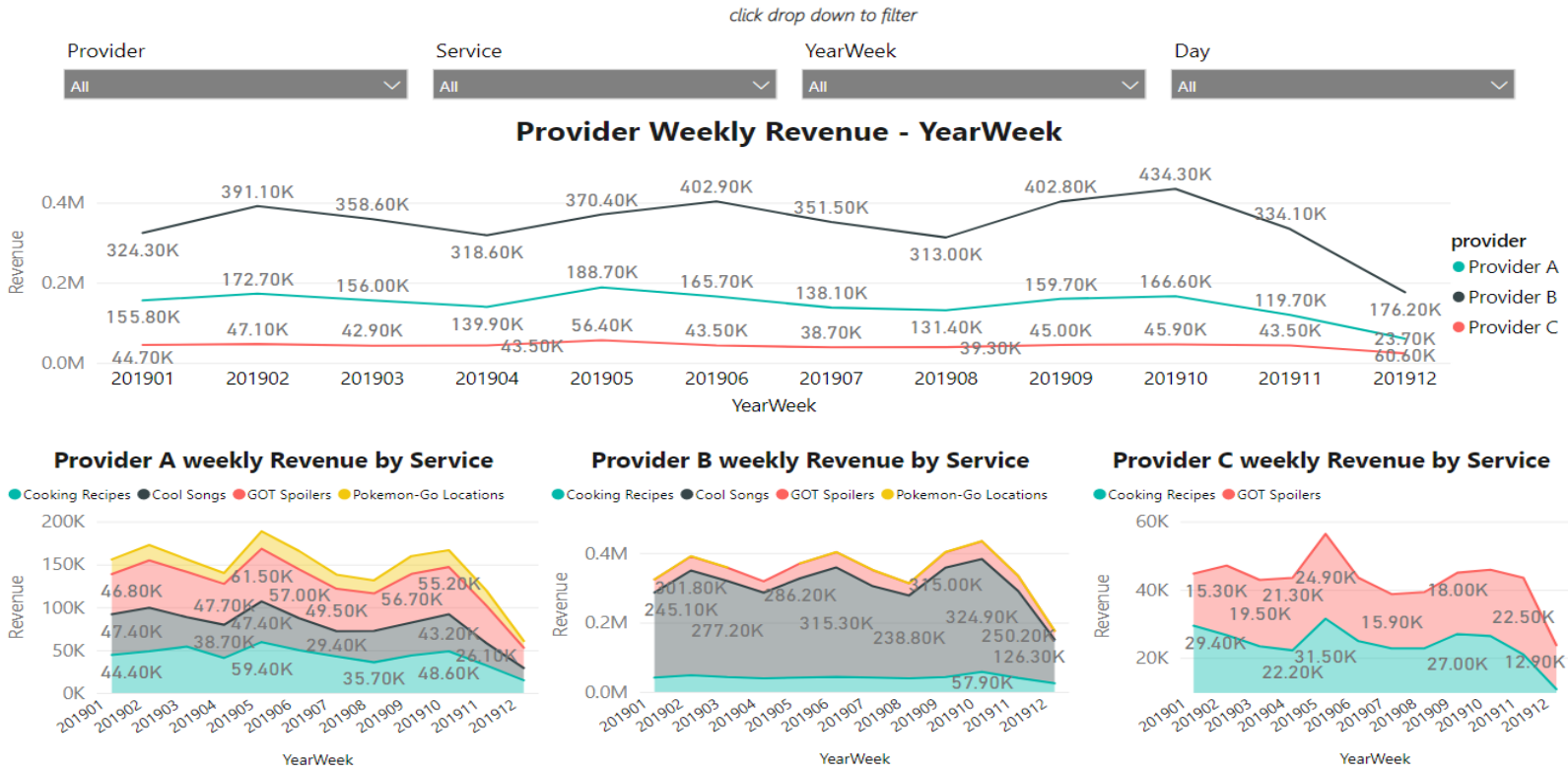
3.)

Yes, 19 unique users have a subscription with both Provider A and B and different services.

19 users	User IDs with different subscriptions across different Providers and/or services							
User_ID	46348431	41977831	32876220	30500718	13879178	13154499	10227890	38952091
	41658851	41249413	30448041	25847480	10110820	5363770	2510874	708441
	32984529	24060999	15495158					

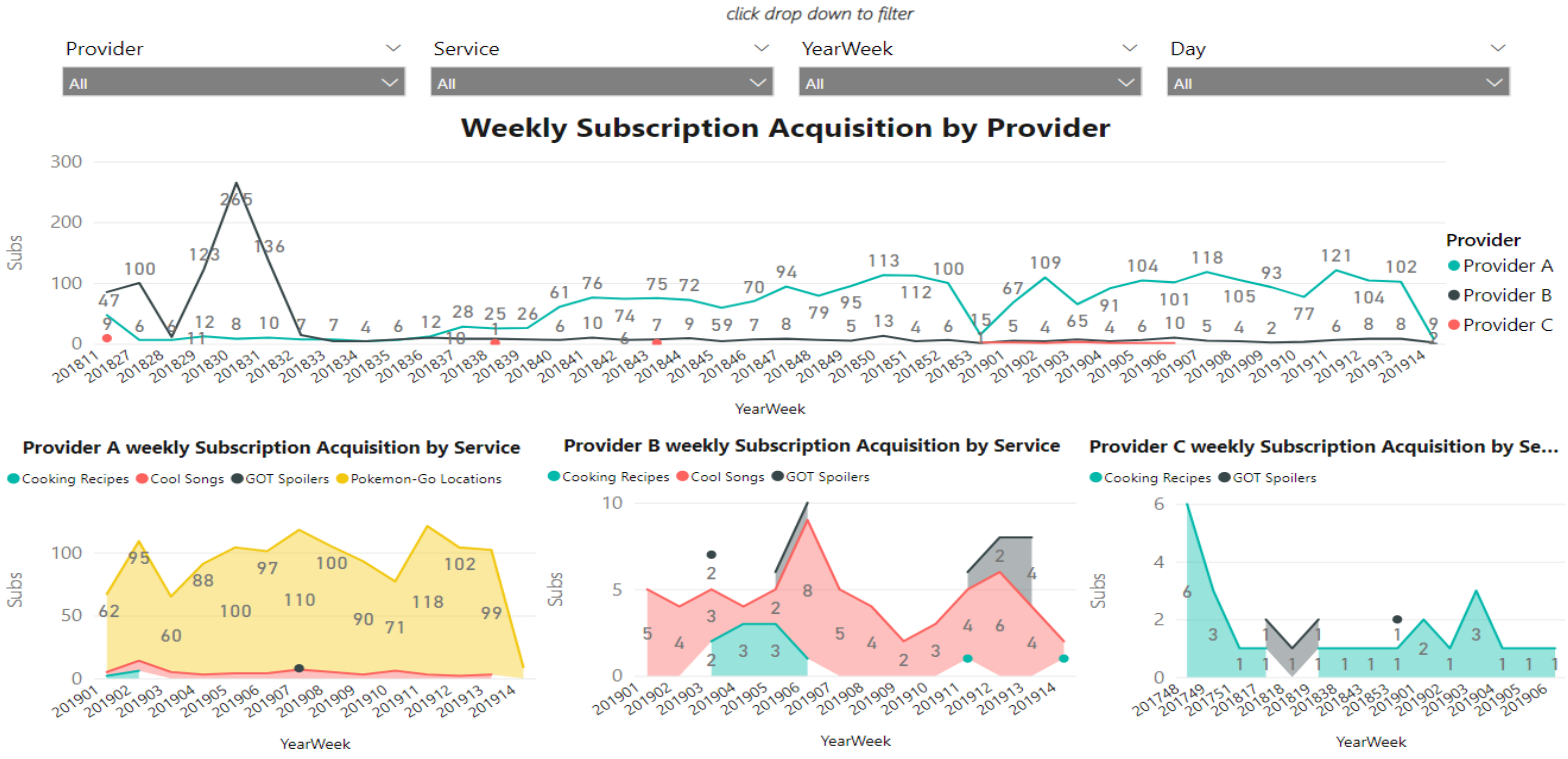
Section B

1.)

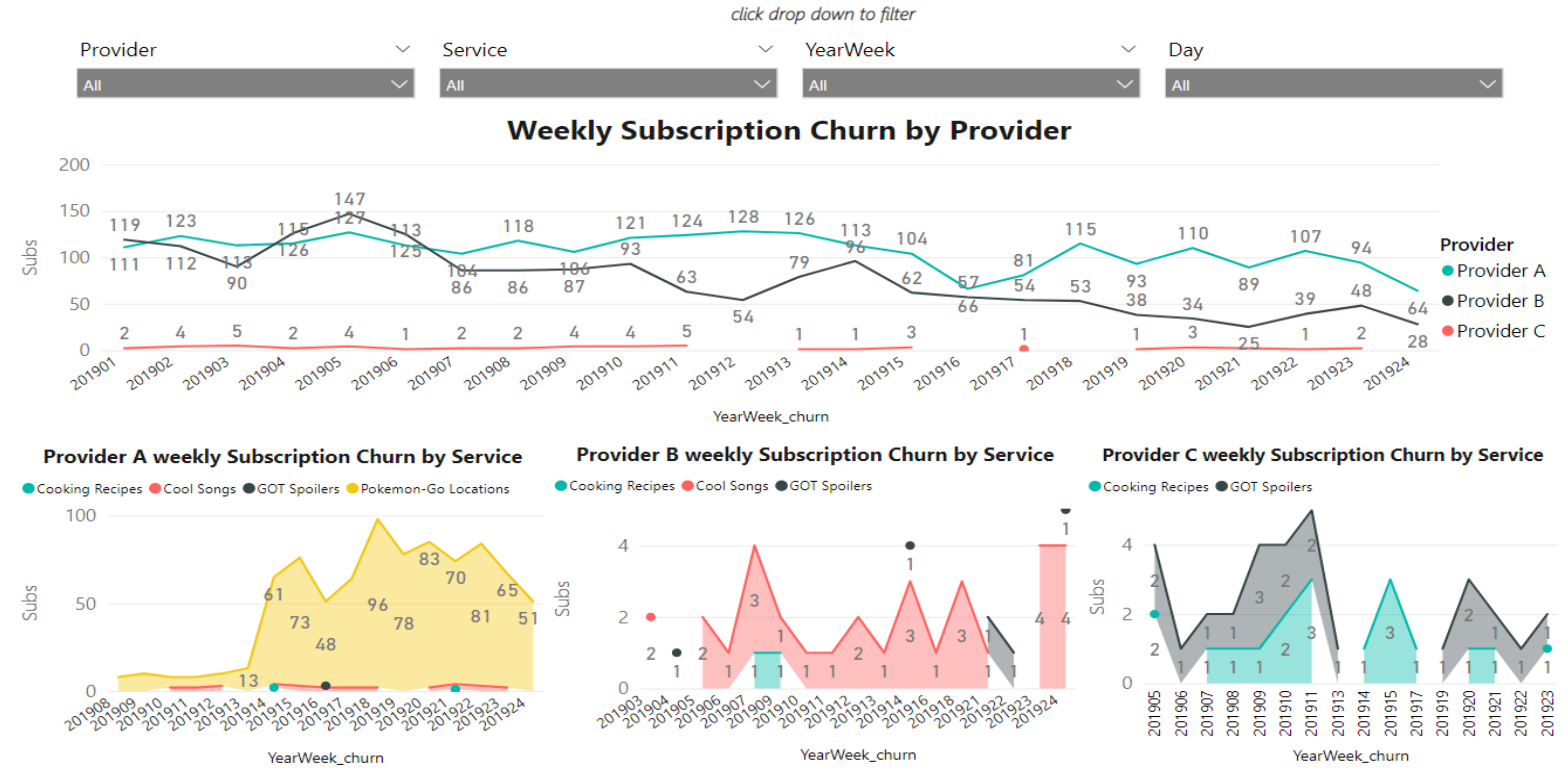


2.)

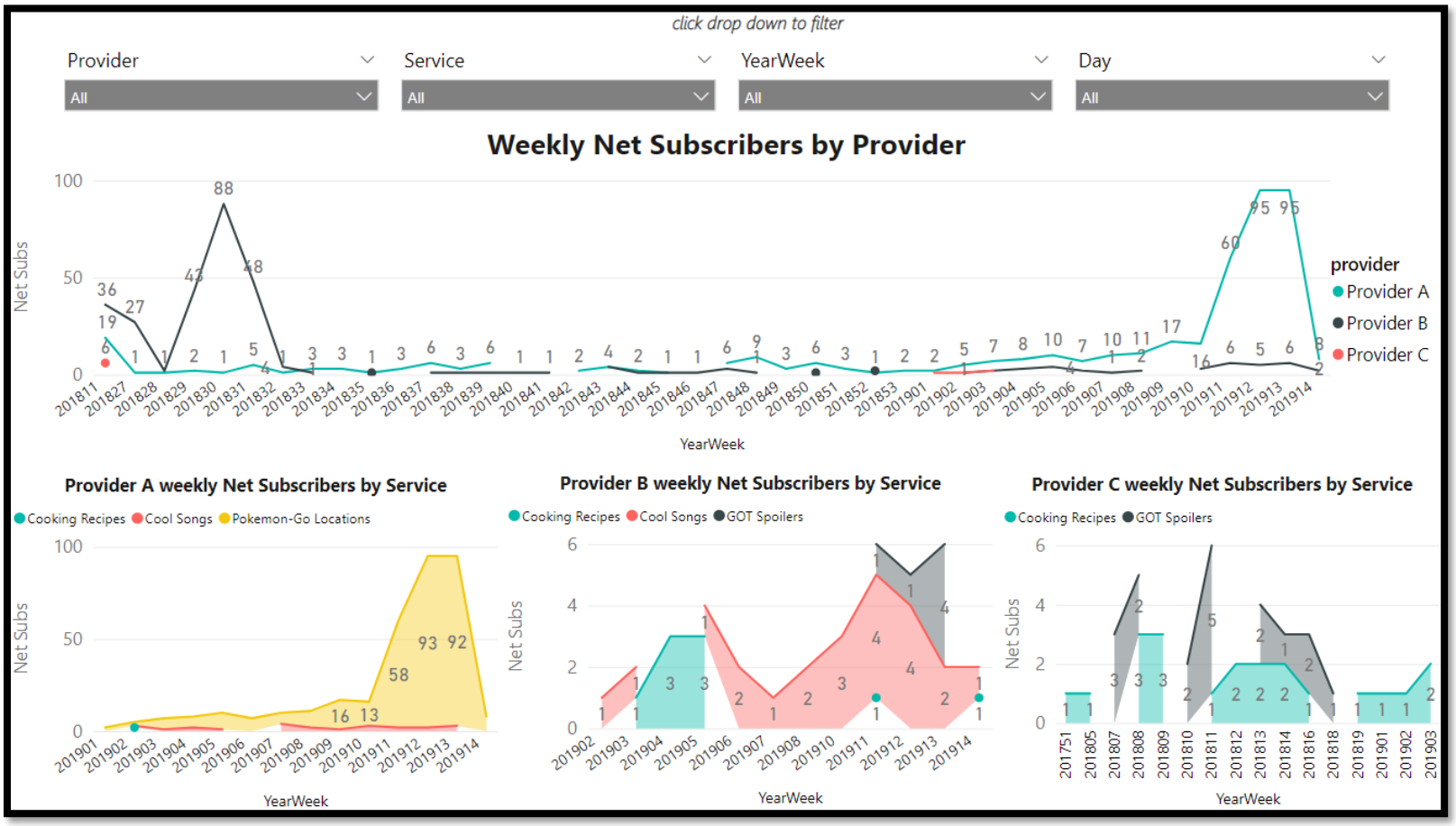
- **Acquisition**



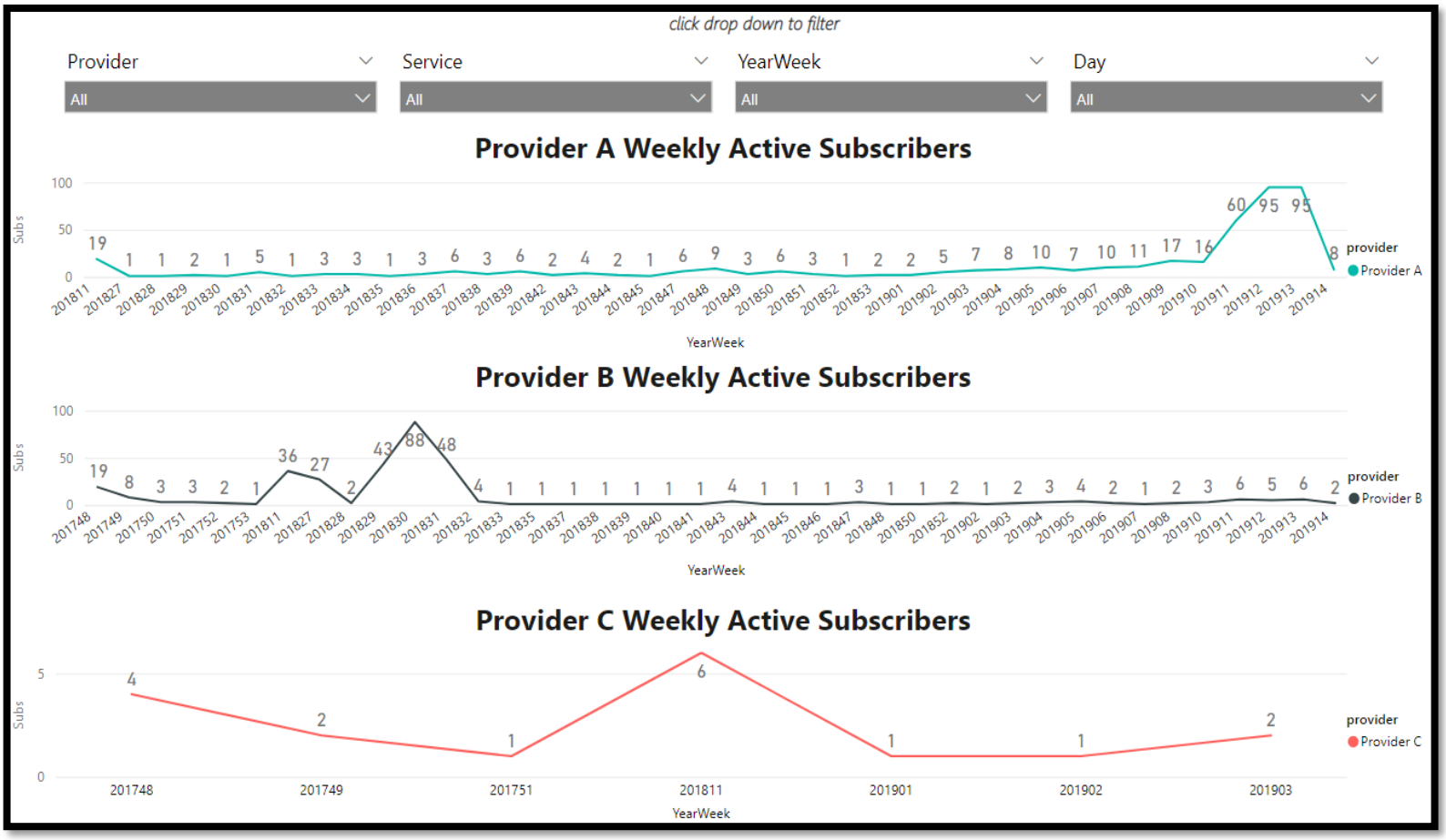
- **Churn**



- **Net Subscribers**



3.)


click drop down to filter
Provider
All
Service
All
YearWeek
All
Day
All
Provider A Weekly Active Subscribers
Subs
YearWeek
provider
Provider A
Provider B Weekly Active Subscribers
Subs
YearWeek
provider
Provider B
Provider C Weekly Active Subscribers
Subs
YearWeek
provider
Provider C

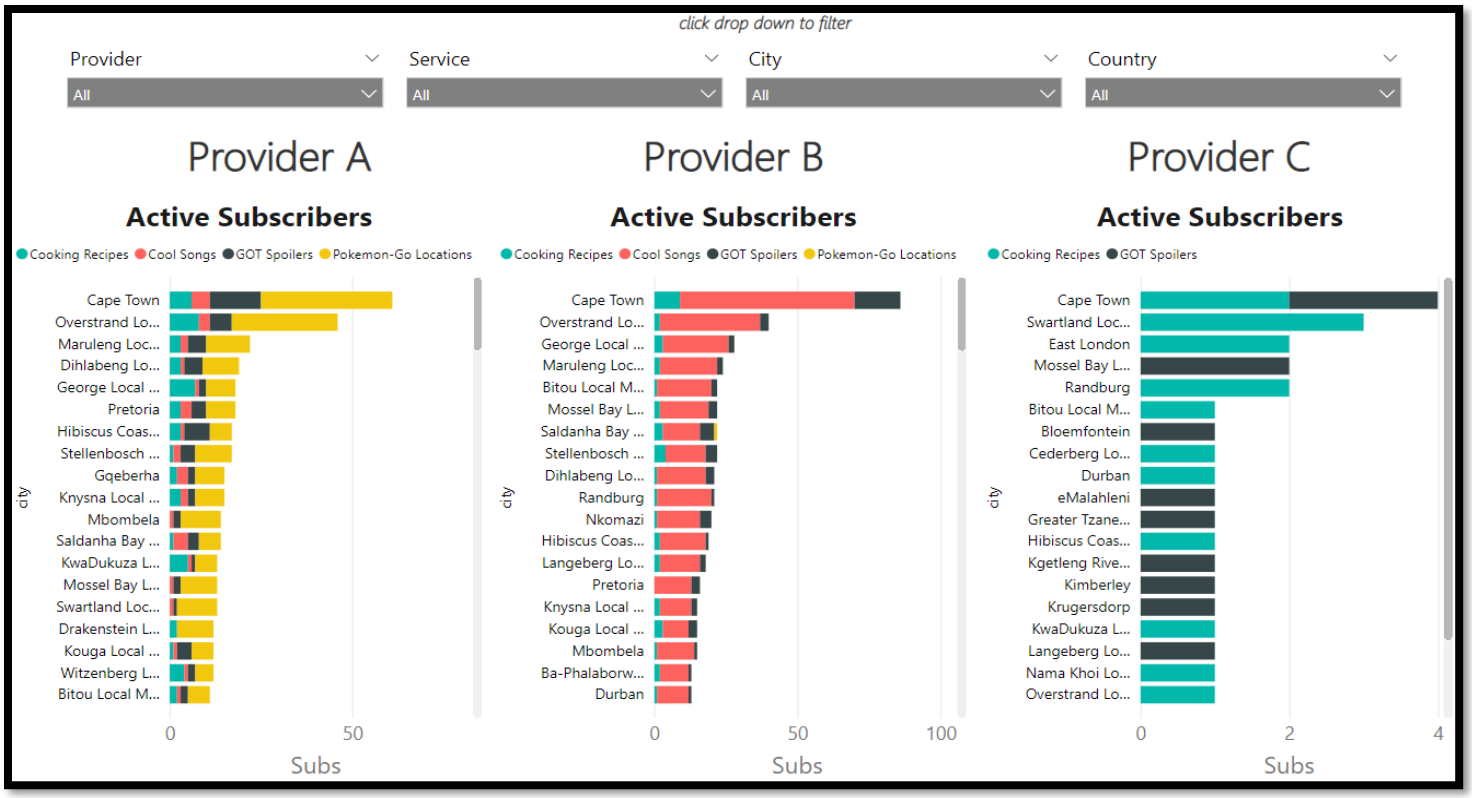
Section C

1.)

– snapshot of the table

user_id	lat	lng	city	country
47324311	-23.95709419	31.16767883	Ba-Phalaborwa Local Municipality	South Africa
46469342	-24.09045029	31.00525284	Ba-Phalaborwa Local Municipality	South Africa
46461990	-24.04663277	31.01746178	Ba-Phalaborwa Local Municipality	South Africa
45994811	-24.04663277	31.01746178	Ba-Phalaborwa Local Municipality	South Africa
43715889	-23.97286987	31.1715107	Ba-Phalaborwa Local Municipality	South Africa
43285256	-24.08176041	31.03813171	Ba-Phalaborwa Local Municipality	South Africa
41946246	-23.94628906	31.14987755	Ba-Phalaborwa Local Municipality	South Africa
40940805	-23.97286987	31.1715107	Ba-Phalaborwa Local Municipality	South Africa
39475828	-24.08176041	31.03813171	Ba-Phalaborwa Local Municipality	South Africa
38227215	-23.95709419	31.16767883	Ba-Phalaborwa Local Municipality	South Africa
35768312	-23.88367462	31.15247726	Ba-Phalaborwa Local Municipality	South Africa
35309922	-23.95709419	31.16767883	Ba-Phalaborwa Local Municipality	South Africa
33851130	-24.04663277	31.01746178	Ba-Phalaborwa Local Municipality	South Africa

2.)



- Cape Town has the most active subscribers across all Providers. In terms of Services, this City has most active subscribers in GOT Spoilers, Pokemon-Go and Cool Songs for

Provider A. While in Provider B, it has most active subscribers in Cool Songs, Cooking Recipes and GOT Spoilers, Provider C it has most active subscribers in GOT Spoilers only.

- Swartland has most active subscribers in Cooking Recipes only for Provider C.

Section D

1.)

Methodology

- I have first created a new variable/column on the Subscription Data file called “Subscription Length”. This is a length/duration (in days) between subscription start date and end date.

	subscription_id	club_id	created	channel	subscription_start	subscription_end	total_billed	billing_rate	billing_cycle	subscription_length_days
0	1	98546	7/16/2020 11:46	WAP	2020-07-16 11:46:00	2020-07-16 11:49:00	0	3	DAILY	0.0
1	2	98546	7/16/2020 15:21	WAP	2020-07-16 15:21:00	2020-07-16 15:28:00	0	3	DAILY	0.0
2	3	98546	7/16/2020 15:35	WAP	2020-07-16 15:35:00	2020-08-13 12:33:00	975	3	DAILY	27.0
3	4	98546	7/17/2020 15:02	WAP	2020-07-17 15:02:00	2020-07-23 10:45:00	300	3	DAILY	5.0
4	5	98546	7/21/2020 12:16	WAP	2020-07-21 12:16:00	2020-07-21 12:28:00	0	3	DAILY	0.0

- I then used the column to check its relationship with total billed column. This provides a clear view on whether a subscriber is billed more the longer they are subscribed to their services. Using the Pearson correlation method, it’s clear that the longer a subscriber is with their services the more they’re billed.

- I then plot the distribution of total billed and subscription length column individually to get their 25^{th}, 50^{th} and 75^{th} percentile.

- **A Quantiles approach** is then used to bin the two columns “Total billed”, and “Length of subscription” based on their 25^{th}, 50^{th} and 75^{th} percentile values. This method groups subscribers into 25% population based on their similarities of Subscription Length and Total Billed.

- I then created a variable called **LS_TB (Lenght_subscription and Total Billed)**. This creates a score of **1-4** for each subscriber **LS_TB** variable, **4 - "Top Subscriber"** being a subscriber who subscribed for longer and billed higher, **1 - "worst subscriber"** - subscribed for few days and billed less. Both Total Billed and Length of Subscription are binned and assigned a score of 1-4 separately, with new variables called TB_Score and LS_Score respectively.

- TB_Score and LS_Score are then combined to form a single LS_TB_score. A Subscriber with **LS_TB Score of 11** would be **"worst subscriber"** since they have subscribed for a shorter period and also billed less. A subscriber with **LS_TB_score of 44** is a Top subscriber since they have subscribed for longer and billed high.

- Once a single **LS_TB_Score** variable is created for each subscriber, I then assigned personalized names to make it easier to understand. These names are assigned based on their LS and TB Scores. The variable with the personalized names is called **LS_TB_Segment.**

Description of the LS_TB Segments

- **Needs Attention** - these are subscribers who were billed high, but they didn't stay long with the services they subscribed to. Finding ways to re-activate these subscribers and have them stay longer could generate more profit to the business.
- **Top Subscriber** - These are subscribers who stay long with the services they subscribed to and are also billed high. These are subscribers who are worth retaining as they're most profitable.
- **Underperforming** - Subscribers who don't spend long with their services and are also billed less.
- **Top Subscriber-Less Bill** - These subscribers stay long with their subscriptions, but they're billed less. It's worth finding strategies to get them subscribe to other content that appetize them to increase revenue.

Snapshot of the resulting table

subscription_id	club_id	created	channel	subscription_start	subscription_end	total_billed	billing_rate	billing_cycle	subscription_length_days	LS_Score	TB_Score	LS_TB_Score	LS_TB_segment
1	98546	7/16/2020 11:46	WAP	7/16/2020 11:46	7/16/2020 11:49	0	3	DAILY	0	1	1	11	Underperfoming
2	98546	7/16/2020 15:21	WAP	7/16/2020 15:21	7/16/2020 15:28	0	3	DAILY	0	1	1	11	Underperfoming
3	98546	7/16/2020 15:35	WAP	7/16/2020 15:35	8/13/2020 12:33	975	3	DAILY	27	2	4	24	Needs Attention
4	98546	7/17/2020 15:02	WAP	7/17/2020 15:02	7/23/2020 10:45	300	3	DAILY	5	2	3	23	Needs Attention
5	98546	7/21/2020 12:16	WAP	7/21/2020 12:16	7/21/2020 12:28	0	3	DAILY	0	1	1	11	Underperfoming
6	98546	7/21/2020 12:26	WAP	7/21/2020 12:26	7/21/2020 12:46	0	3	DAILY	0	1	1	11	Underperfoming
7	98546	7/21/2020 12:33	WAP	7/21/2020 12:33	7/21/2020 12:38	0	3	DAILY	0	1	1	11	Underperfoming
8	98546	7/22/2020 15:13	WAP	7/22/2020 15:13	8/13/2020 13:37	0	3	DAILY	21	2	1	21	Underperfoming
9	98546	7/22/2020 15:14	WAP	7/22/2020 15:14	7/22/2020 15:16	0	3	DAILY	0	1	1	11	Underperfoming
10	98546	7/22/2020 15:16	WAP	7/22/2020 15:16	7/22/2020 15:18	0	3	DAILY	0	1	1	11	Underperfoming
11	98546	7/22/2020 15:19	WAP	7/22/2020 15:19	7/22/2020 15:29	0	3	DAILY	0	1	1	11	Underperfoming
12	98546	7/22/2020 15:39	WAP	7/22/2020 15:39	8/13/2020 13:56	3000	3	DAILY	21	2	4	24	Needs Attention

Summary of the LS_TB_segments

LS_TB_segment	total_billed mean	subscription_length_days mean	count_of_subscriptions
Needs Attention	1084.4	16.5	34951
Top Subscriber	1897	63.6	88278
Top Subscriber-Less Bill	0	60.9	76143
Underperfoming	0	6.7	130117

Observations

- The **Top Subscriber** are billed more than other subscribers **(average 1897 billed)** and they also spend longer on their subscriptions **(average 63 days)**
- The **"Needs Attention segment"** are the second most billed but they do not spend long on their subscription (only average 16 days).
- **Top Subscriber-Less Bill** - billed less but stay long with their subscriptions.
- **Underperforming** - least billed and low duration on their subscriptions.

***NB**: Resulting table (xlsx file) and code shared on python/Jupyter notebook file.*

----------END!!! ----------