

ITV Step Up Data Challenge

Claire Jeynes, Georgie Howell, Ashild Kummen



Welcome to the 'Step Up' ITV live online career challenge

What will you need?

- A good wifi connection
- A notebook and pen
- A quiet environment where you will not be disturbed for 45 minutes

What do we expect?

- Respectful and appropriate behaviour throughout this 45 minute briefing session
- A positive attitude and willingness to learn
- There will be a number of opportunities throughout today's presentation to ask question using the chat function

How will you benefit from completing this challenge?

- The opportunity to solve ITV's business problems
- To learn about how your curriculum learning at school is used in the 'real world'
- Understand how valuable 'YOUR' 'key skills' are when solving this problem and delivering the solution
- Gain a virtual work experience certificate

Any technical issues during this session - please email info@digdata.online



Meet the team



Claire Jeynes

Lead Data
Analyst



Georgie Howell

Data Analyst



Åshild Kummen

Machine
Learning
Engineer

ICEBREAKER CHALLENGE

Do you watch ITV?



itvX

What year was the online streaming platform (iTV Hub initially) launched?

A

1995

B

2003

C

2008

What year was the online streaming platform (iTV Hub initially) launched?

A

1995

B

2003

C

2008

What was the most viewed programme in 2022 on ITV Hub / ITVX?

A



B



**FIFA WORLD CUP
Qatar2022**

C

this
morning

What was the most viewed programme in 2022 on ITV Hub / ITVX?

A



B



**FIFA WORLD CUP
Qatar2022**

C

this
morning

How many customers on average do we have viewing on ITVX per month?

A

B

C

10

12

15

million

million

million

How many customers on average do we have viewing on ITVX per month?

A
10
million

B

12

million

C

15

million

ITV

Who are we?

ITV Structure



*ITV
Studios*

One of the biggest global creators, producers and distributors in the world

We create, produce, and distribute a broad range of programmes, including drama, entertainment and factual. Our customer base is diverse, producing for international television broadcasters and global streaming platforms.



*Media &
Entertainment ***

(M&E) is the home of ITV's family of channels and platforms – the largest family of free-to-air commercial channels in the UK, with programming delivered across multiple platforms.

M&E is made up of two parts –
Streaming (ITVX) and Broadcast.

What do ITV do?

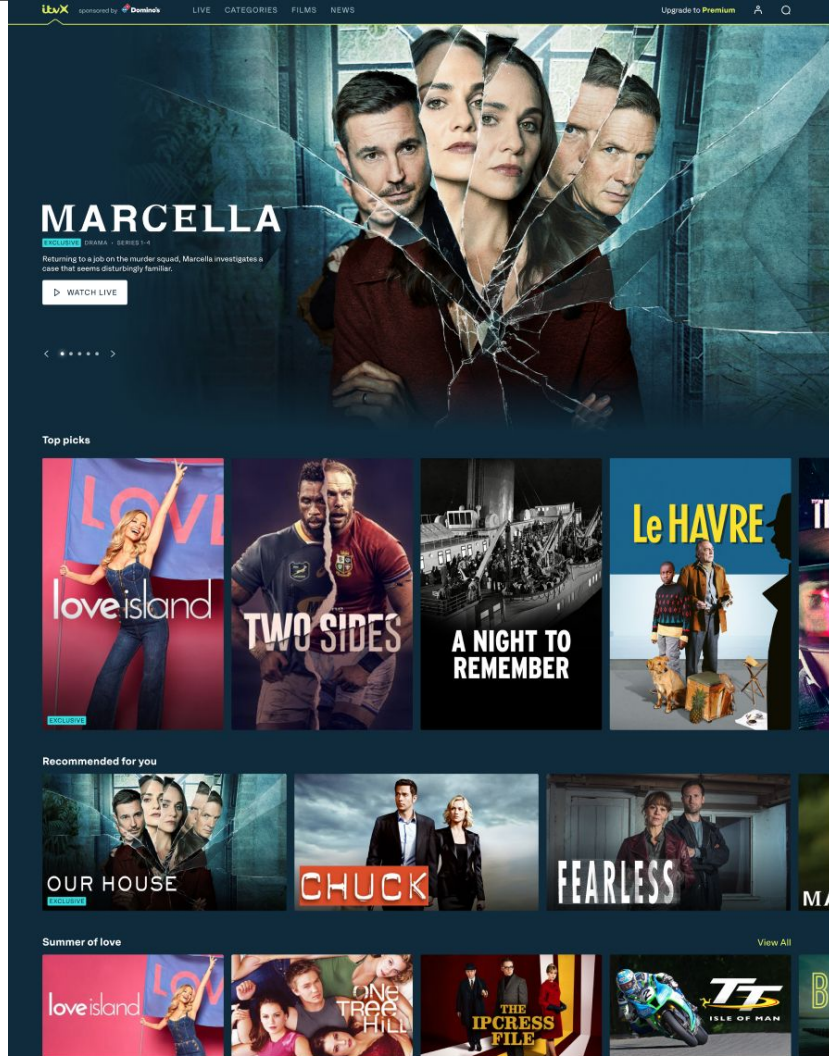
- ITV is a British TV powerhouse that brings all the best shows to your screens!
- We've got a whole bunch of channels like the main ITV, ITV2, ITV3, ITV4, and CITV.
- We've got everything from juicy dramas, to belly laughing comedies, to nail-biting game shows, and reality TV that'll have you on the edge of your seat!
- ITVX IS the new streaming service - where you can watch live and on-demand content online.
- We're not just big in the UK, We're also making waves internationally with production companies all over the world.
- Oh, and we make money through ads and sponsorships.



What is ITVX?

At **ITV**, we love creating some of the nation's best-loved and most-watched programmes. We connect with millions of people every day, make programmes the public can't get enough of and reflect and shape the world we live in... and we do all this through the power of creativity!

ITV On Demand is the home of the platform that powers the recently launched ITV X. In 2021 ITV served over 1.8 billion videos to 28 million users across 29 platforms including Amazon Fire, Xbox, Android, and iOS devices to name but a few.



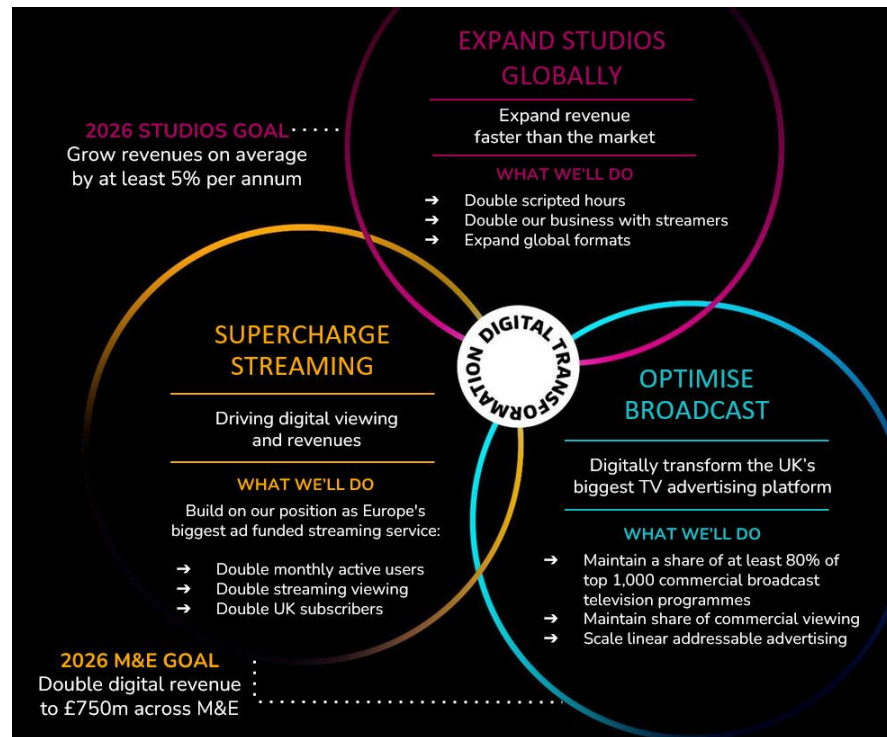
What data do we collect?

- Viewing data
- Customer Information
- Interactive (voting data)
- Subscription information

This is just a small snapshot of the huge wealth of data ITV work with

What do we use data for?

- Reporting on business performance
- Generating insight to inform decision making that improves the customer experience/interactions with ITV



End result...

itvX CORE KPIs

a view of user engagement with ITV's flagship digital proposition & how this ladders up to total streaming hours

Full Month: Dec 2022

1 Dec 2022 - 31 Dec 2022

CORE KPI TREE

user lifecycle stage

total streaming hours

consumption hours

10.1M

↑ 67.9%
versus last year

viewers

maus watched >2 mins content

4.0M

↑ 93.6%
versus last year

streaming hours per viewer

average streaming hours per viewer

3.5

↓ -13.2%
versus last year

users

monthly active users

200K

↑ 76.6%
versus last year

conversion to view

% that streamed >2mins content

76%

↑ 9.6%
versus last year

breadth of interest

programmes streamed

1.1

↓ -23.4%
versus last year

depth of viewing

streaming mins per programme

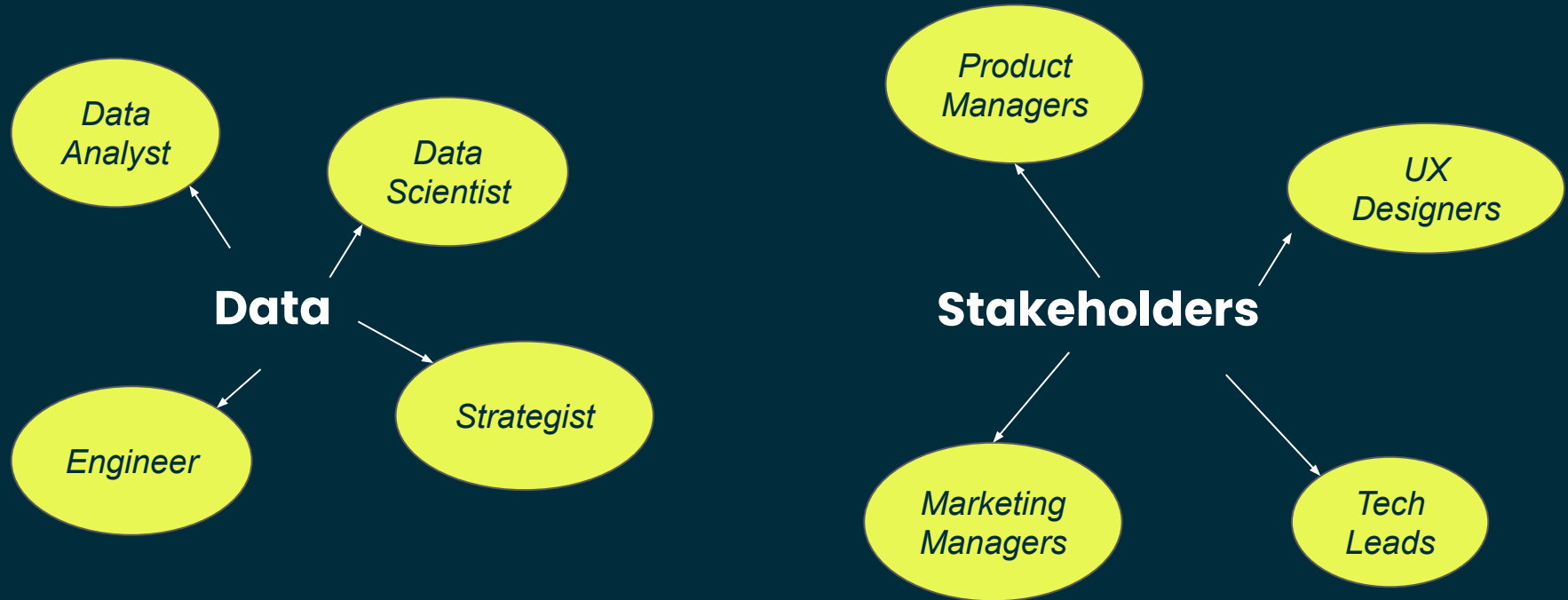
126

↑ 13.3%
versus last year



*fake data

Data Roles and stakeholders



What tools do we use?



databricks



Google Cloud Platform



The Challenge

How do ITV promote content on their new streaming platform ITVX effectively to their viewers?

Overview:

ITV have just launched their new streaming platform ITVX and have commissioned new shows by partnering with third parties such as Warner brothers to expand the breadth of ITVX content and the number of hours available to watch (from 3,000 to 15,000!).

Step Up Challenge:

With an average of 30 million registered users, how can we ensure that they discover this new great content, increase the number of hours they stay on the platform and are aware of the new features that ITVX has to offer?



How do ITV promote content on their new streaming platform ITVX effectively to their viewers?

ITVX is launching a new original show every week, 'A Spy Among Friends' is one of the first original shows on ITVX in 2022

- **Task 1 - Data Strategy**

- We want to know what the customer base viewing habits look like?

- **Task 2 - Data Science**

- Who should we be promoting this show to?

- Based on what you have learnt, what recommendations would you give to product and /or marketing to promote 'A Spy Among Friends'.



Data Strategy Task

Task 1: Data Strategy

We want to know what the customer base viewing habits look like?

- Please profile the customers in some suggested groupings below so we can understand our customers viewing behaviours better. Knowing how customers behave helps us inform on decisions such as marketing comms on who we should target to watch the new originals for ITVX
 - *Genre, Breadth of Interest, Consumption*



Data

Each registered user has their own unique **user_id**

The duration of the entire streaming session in seconds
Remember, multiple programmes can be watched within this session

The total number of seconds the piece of content was watched for

Each episode of a programme has a unique episode_id

user_id	stream_id	platform	session_duration_seconds	is_weekend	session_start_datetime	session_end_datetime	stream_type	consumption_seconds	programme_id	episode_id	programme_title
camyxhGynh+mmQ6aC	d38clC83-9341-417b-bae7-eac	Mobile	186	FALSE	11/01/2022 06:56:48	11/01/2022 06:59:54	vod	185.938	1/1096	780383	MIDSOMER MURDERS
Ql9ga/UlVRWfknjVCgId	3D585E03-ACF2-4CD8-905D-0	TV	276	FALSE	11/01/2022 07:02:43	11/01/2022 07:07:19	vod	276.021	1/0694	1/0694/10783	CORONATION STREET
Ql9ga/UlVRWfknjVCgId	9AE5188F-B31F-4FFF-ADCC-8C	TV	1185	FALSE	11/01/2022 07:08:07	11/01/2022 07:27:52	vod	1185.179	1/0524	1/8694/9507	EMMERDALE
zXTs6xNqnxF7Q7g0pTSjt	23d6ee1b-12f7-4b78-a104-1896	TV	585	FALSE	11/01/2022 07:41:57	11/01/2022 07:51:42	vod	584.84	2/7556	2/7556/0005	THE WALK-IN
ckfiYDo/UFRDFhq2H5bX	A4D68058-7236-46DE-AA73-3f	TV	3357	FALSE	11/01/2022 08:18:57	11/01/2022 09:14:54	simulcast	843.5107096	1/9360	1/9360/3221	LORRAINE
m5LUnpWBZzhZDQ0Vsr	Id6640e4-e529-4e67-b40e-07	TV	1218	FALSE	11/01/2022 08:22:52	11/01/2022 08:43:10	vod	1218	1/0694	1/0694/10783	CORONATION STREET
CvRvICNLhZBFdbxrmOf	5f6e5d6b-fab6-420e-9bb9-R2	TV	11138	FALSE	11/01/2022 08:33:02	11/01/2022 11:38:40	simulcast	3231	1/1960	2/6292/0963	THIS MORNING
7r/kxWUPCU9/EQ9H+K	14a80b69-c878-4659-9e65-1b	Desktop	0	FALSE	11/01/2022 08:33:18	11/01/2022 08:33:18	vod	0	2/3398	2/3398/0160	THE REAL HOUSEWIVES OF CHE

Each viewing session has a unique **stream_id**. Multiple programmes can be watched in one viewing session

VOD - Video On Demand/Catch Up
Simulcast - Live TV

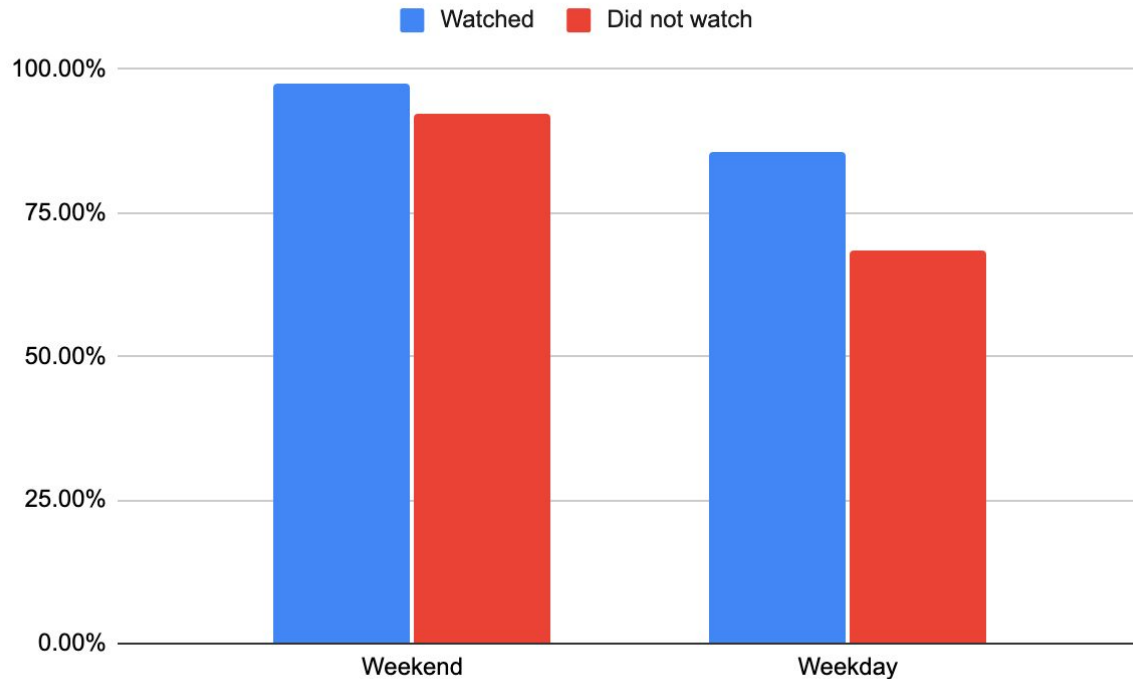
Each programme has a unique programme_id
E.g. Coronation Street's programme_id is 1/0694

**see how to guide for full dictionary*

Hints and Tips

- Feel free to use whatever methodology you want to summarise the data, this can be completed in excel.
 - For those with sql knowledge, feel free to use sql to analyse customer behaviour.
- Focus on using the flag to compare customers who have viewed *A Spy Among Friends* vs who haven't. What behavioural traits can you decipher that are different.
- Feel free to visualise your findings in any format you prefer, this can be in excel, a slide in powerpoint or word document. Be as creative as you want.
- Pivot table, line graphs and bar charts could be your best friend!

EXAMPLE : From the graph below, we can see the variance between the target audiences vs non target is higher during the week, we could suggest to marketing to focus comms on mid week to influence mid week viewing adoption



Data Science Task

Task 2: Data Science

Use the data to design a model which predicts the likelihood of a user watching the new ITVX exclusive 'Spy Among Friends'.

- We will provide you with a notebook and some examples to play around with (Python, Jupyter Notebook)
- If you already have experience and feel comfortable taking your own approach - we recommend you start from scratch in whatever environment you prefer.
- Create at least three models with different feature sets
- Evaluate the performance of the models and the pros and cons of each
- Interpret the models and provide recommendations for the business on next steps: use visualisations and provide key model metrics



Data

Each row represents
one unique user id

The top 3 programmes and genres
that user has watched

The total number of seconds a
user watched something on ITVX
on a *weekend*

The total number of seconds a
user watched something on
ITVX overall

user_id	n_sessions	n_programmes_watched	top_3_programmes	top_3_genres	total_weekend_consumption_seconds	total_consumption_seconds
+1FV4TibF3C+	13	5	['WITHOUT SIN', 'EMMERDALE', 'CORONATION STREET']	['DRAMA']	1561.353	10650.877
+AprntNc0A0Yl	26	7	['WORLD CUP', 'I'M A CELEBRITY... GET ME OUT OF HERE!', 'LOVE ISLAND']	['ENTERTAINMENT', 'SPORT', 'DRAMA']	7201.881194	52660.12492083001
+B/+uicTINNsFS	39	8	['A SPY AMONG FRIENDS', 'A CONFESSION', 'WITHOUT SIN']	['DRAMA', 'ENTERTAINMENT', 'OTHER']	0.0	78485.44204222999
+BpTxKOa0+lzi	67	18	['WORLD CUP', 'FEARLESS', 'THE WALK-IN']	['SPORT', 'ENTERTAINMENT', 'DRAMA']	70648.5916638	176989.16734430892
+DYiONWpxJsu	231	40	['AGATHA CHRISTIE'S POIROT', 'LEWIS', 'MIDSOMER MURDERS']	['DRAMA', 'OTHER', 'FACTUAL']	68043.45445744101	188989.3533681411
+Ni+NSzcfeaq/r	7	7	['A SPY AMONG FRIENDS', 'DOC MARTIN CHRISTMAS SPECIAL 2022', 'MATER']	['DRAMA', 'FACTUAL']	3457.531	15770.636999999999
+PDN5TuwsCX	34	6	['A SPY AMONG FRIENDS', 'STONEHOUSE', 'KAREN PIRIE']	['DRAMA', 'OTHER']	6708.169	37421.959
+Qg/adlna+sibz	4	1	['A SPY AMONG FRIENDS']	['DRAMA']	0.0	1583.265

Summary metrics per user

... and 31 other columns!

Data Dictionary :

Data Science : The DATA

1763 Rows, 31 columns

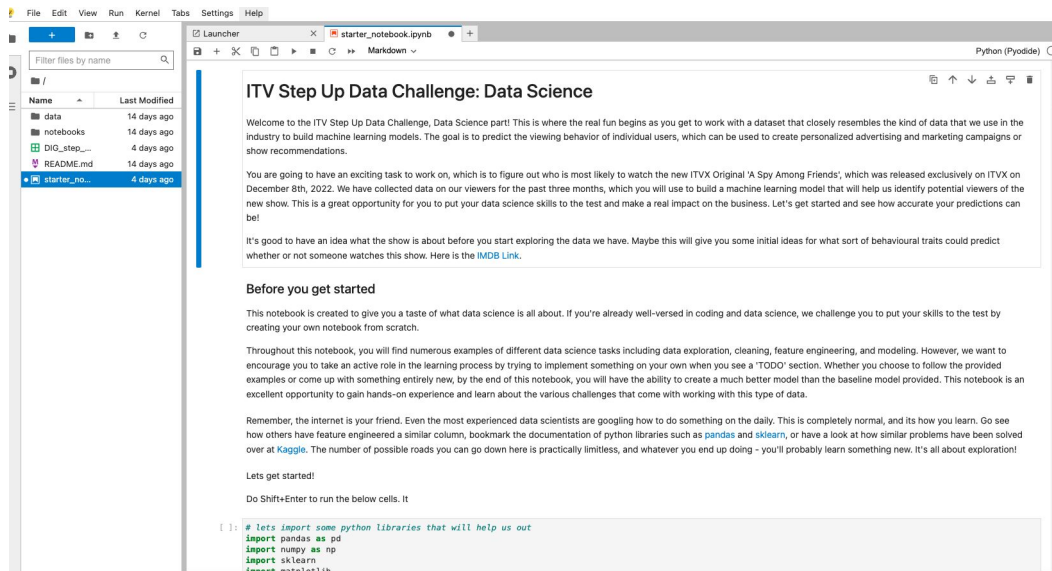
The '*any_spy_among_friends_consumption*' column is our target, showing whether a customer watched 'A Spy Among Friends'. Watched = TRUE, Did not watch = FALSE.

Disclaimer : This is not real ITV data but a representation of the types of data we do use everyday

Data Field	Data Explanation
user_id	All data grouped at a distinct user ID level, user_ID unique to a specific customer
n_sessions	Total number of distinct sessions
n_programmes_watched	Distinct number of programmes a customer has watched any length of content on
n_episodes_watched	Distinct number of episodes a customer has watched any length of content on, note you can have 10 episodes to 1 programme
top_3_programmes	Ordered by consumption, top 3 programmes a customer has viewed
top_3_genres	Ordered by consumption, top 3 genres a customer has viewed

**see how to
guide for full
dictionary*

Starter Notebook



**see how to
guide for full
instructions*



Hints and Tips

- Don't be intimidated by a lack of prior experience! Embrace the opportunity to learn and grow as you work with the data.
- Stay curious and question the data you're working with. If something doesn't look quite right, don't hesitate to investigate further.
- Gather as much information as possible from the data to inform your analysis and modeling.
- Carefully consider which features to include in your model, and how they may impact performance. Keep in mind the goal of adding value to the analysis.
- Remember that you're not alone in this process! The internet is a valuable resource for finding answers and learning new skills. Don't be afraid to seek out help when needed.

Thank You!
Good Luck!