

hybrid

CODE & MVP



Lukáš Hnilička

Nathaly Toledo
Zuzana Jankovská
Garry Putranto

WHY	03
NABC	05
Background	08
Technology & Results	09
Process used	16
Reflection	17
Accomplished - Next	18
Q&A	19

CONTENTS

//02

THE COMPANY BACKGROUND

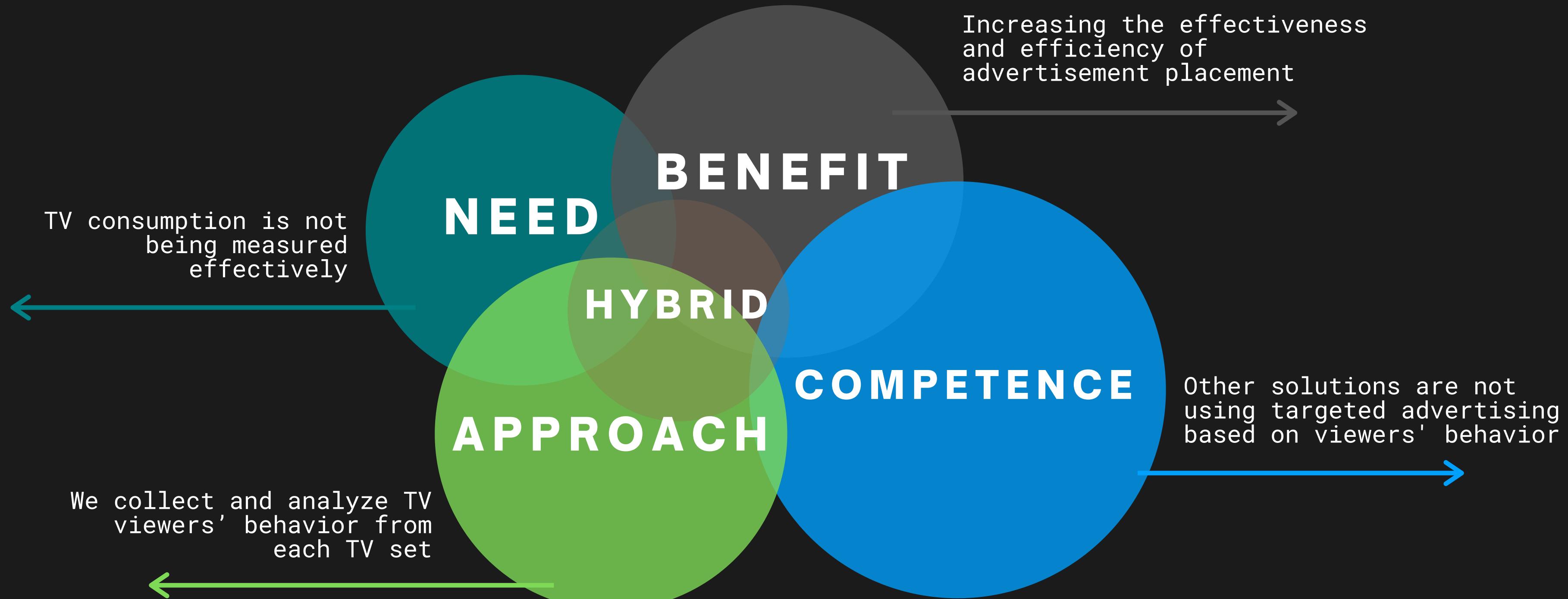
MediaTech Startup

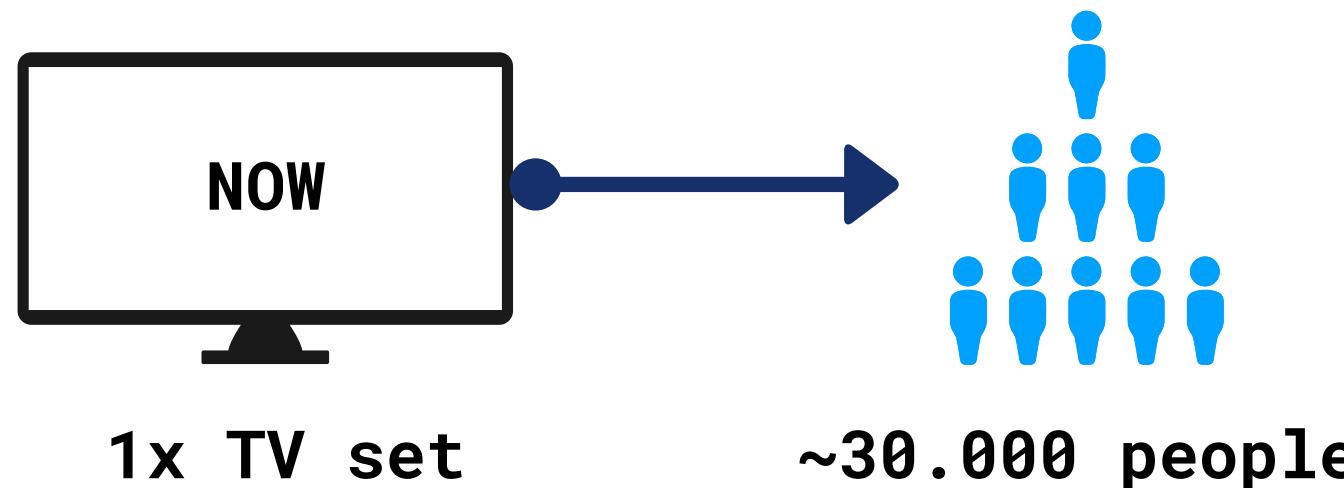
Client: TV broadcasters and TV operators.

Location: Prague, the Czech Republic.

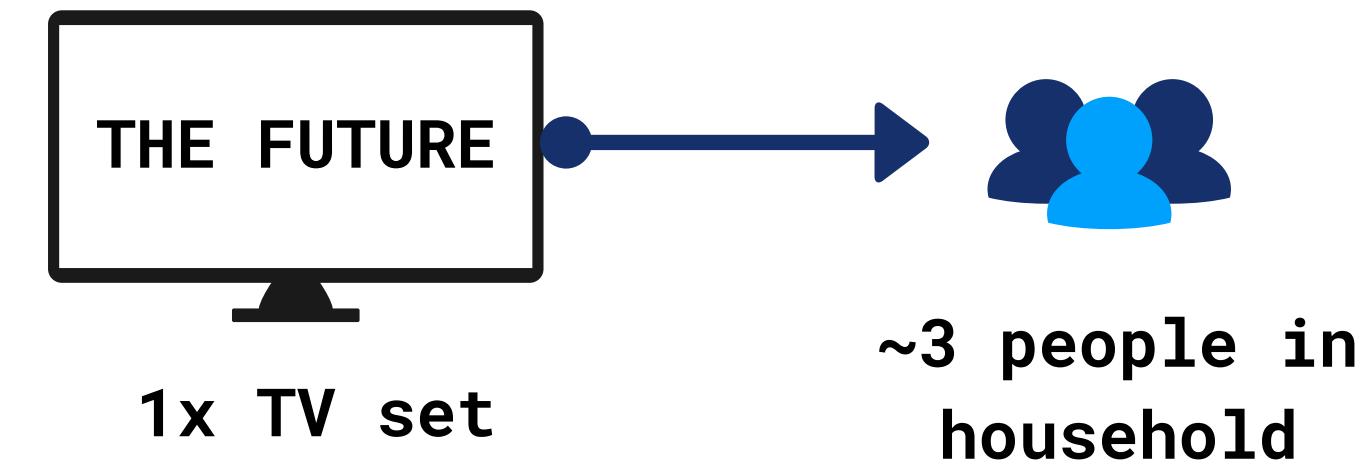
Mission: to develop next-gen TV audience viewership.

VALUE – NABC REMINDER



**FACT #1**

Current measuring technology is inaccurate. Using sampling method when 1 person represents ~30.000 people.

**BENEFIT #1**

A new source of own online data independent to Facebook and Google is created.

FACT #2

Current technology is 20+ years old.

BENEFIT #2

By targeting the right users (potential customers), advertisers can save money and provide additional value for TV viewers

FACT #3

Advertisers are heavily data oriented and measurement is not.

BENEFIT #3

A **customized experience** increases the value of streamed content and personalizes advertisement placement.

WHY US, WHAT FOR?

STATUS QUO



OUR PROJECT



BRIGHT FUTURE



Limited screening of TV viewership, not even Facebook or Google has those data...

Data transformation,
business intelligence
machine learning

We help TV broadcasters and advertisers fulfill business needs while optimizing viewers' experience

...BUT WE DO HAVE THE DATA AND WE KNOW HOW TO GAIN
VALUABLE INSIGHTS ABOUT VIEWER BEHAVIOR AND HELP TV
BROADCASTERS AND ADVERTISERS TO ACT UPON THEM

BACKGROUND - START : END GOAL

//07



RAW DATA

- Channels
- User devices
- Sessions
- Playout / Content displayed



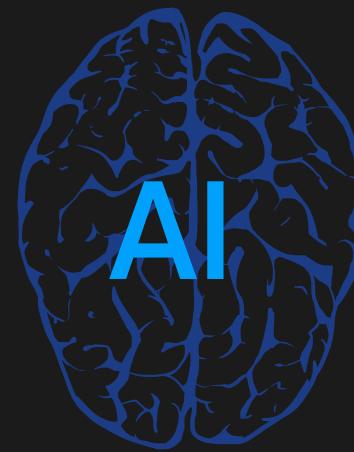
DATA ARCHITECTURE

1. Transformation - Extract, Transfer, Load (ETL)
2. DWH - data modeling
3. Data Storing
4. Build data market place



BUSINESS INTELLIGENCE

1. Implementation BI tool (Tableau)
2. Market place connection (DWH->Tableau)
3. Visualization of case study
4. Visualization of ML forecast



MACHINE LEARNING

1. Definition of use ML for project - focus, target
2. Data Source preparation
3. Inputs - assignment - output
4. Storing to DWH

0



1



2



3



PIONEER

FIRST attempt in the world
RECREATE the TV world

UNIQUE

measure and deliver
personalized
content with
**BIG DATA, DATA
SCIENCE & MACHINE
LEARNING**

brand new concept

MINIMUM VIABLE
PRODUCT created in
less than a week

whole journey from
raw data to machine
learning

AMBITIOUS

B2B, value for the
whole TV market

VALUABLE

TV broadcasters,
advertisers,
TV viewers

TECHNOLOGY



DATA ARCHITECTURE

WHAT WE WERE USING

TECHNOLOGY
- I

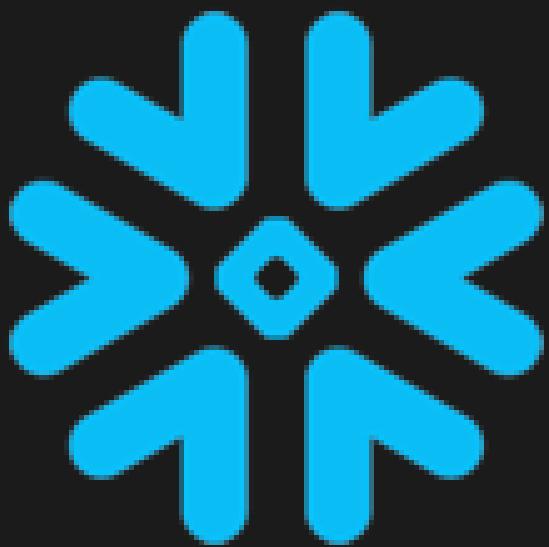


INFLUXDB
TIME SERIES DATABASE

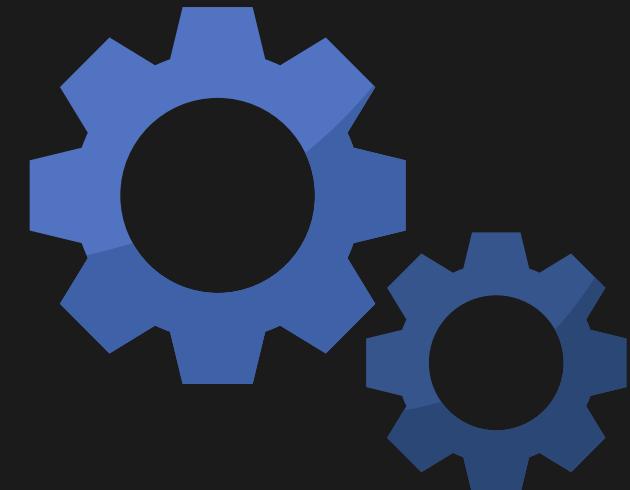
PYTHON



SNOWFLAKE
DATA WAREHOUSE



CREATE
TRANSFORMATION



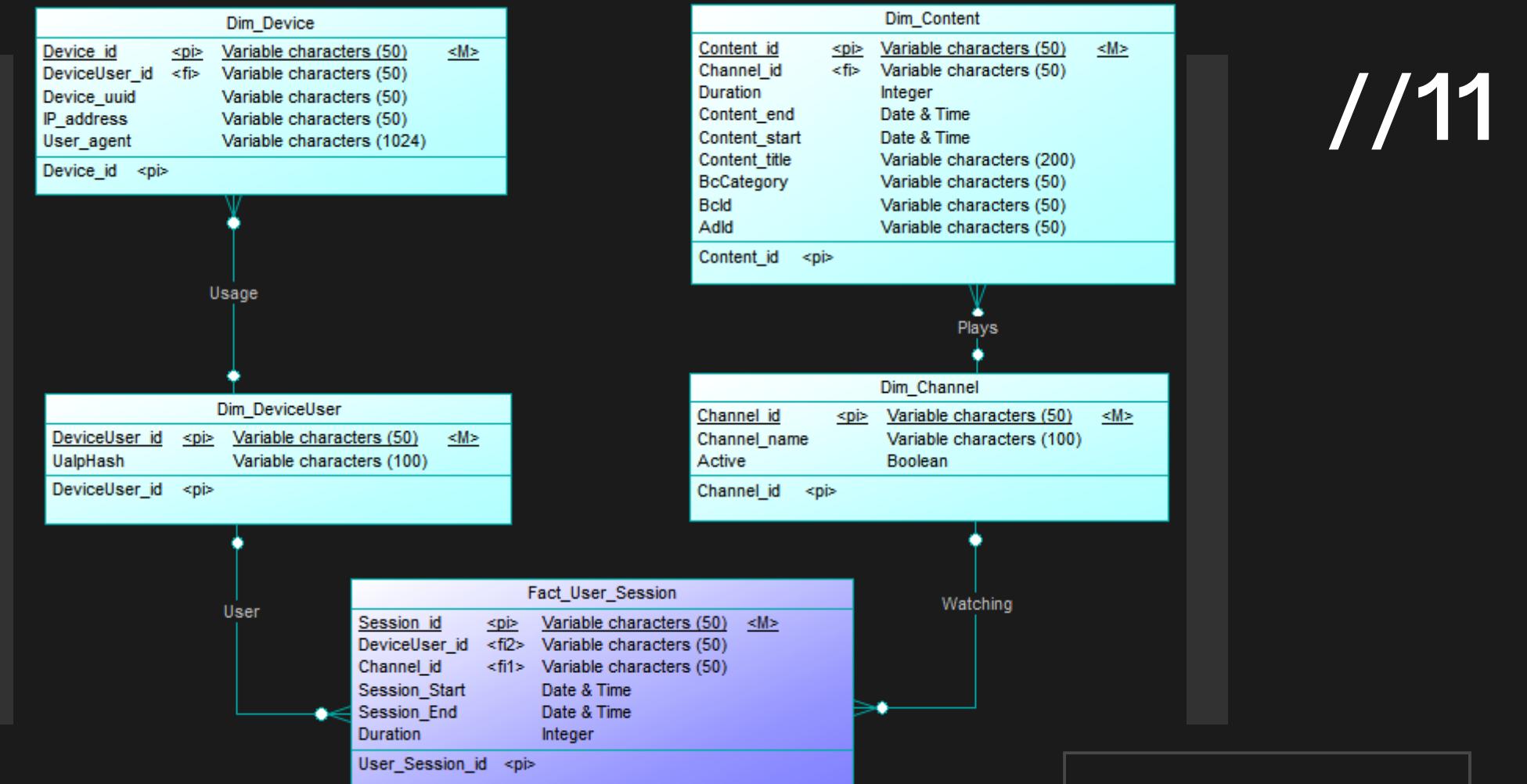
Transferring more than 4 million records
from time stream database to data
warehouse

DATA ARCHITECTURE

Deliverables:

- ETL - Extract, Transform, Load
- DWH Content Viewership
- Snowflake Afterspot SQL
- DWH Logical Data Model

RESULTS - I



RESULT #1 DATA WAREHOUSE ARCHITECTURE

```
-- CREATING DEFAULT OBJECTS
CREATE ROLE AFTERSPOT_ETL_ROLE;
CREATE USER ETL PASSWORD = '*****' DEFAULT ROLE = "AFTERSPOT_ETL_ROLE";
CREATE USER BI PASSWORD = '*****' DEFAULT ROLE = "AFTERSPOT_ETL_ROLE";
CREATE WAREHOUSE AFTERSPOT_WH WITH WAREHOUSE_SIZE = 'XSMALL' WAREHOUSE_TYPE = 'STANDARD' AUTO_SUSPEND = 60 AUTO_RESUME = TRUE;
CREATE DATABASE AFTERSPOT_DB;
CREATE SCHEMA AFTERSPOT_DB.RAW_SCHEMA;
CREATE SCHEMA AFTERSPOT_DB.MARKETPLACES_SCHEMA;

-- SESSION SETTINGS
ALTER USER ETL SET QUOTED_IDENTIFIERS_IGNORE_CASE=TRUE;
ALTER USER BI SET QUOTED_IDENTIFIERS_IGNORE_CASE=TRUE;

-- GRANTS
GRANT ROLE AFTERSPOT_ETL_ROLE TO USER ETL;
GRANT MODIFY, MONITOR, USAGE, CREATE SCHEMA ON DATABASE AFTERSPOT_DB TO AFTERSPOT_ETL_ROLE;
GRANT USAGE ON ALL SCHEMAS IN DATABASE AFTERSPOT_DB TO ROLE AFTERSPOT_ETL_ROLE;

GRANT ALL ON ALL TABLES IN SCHEMA MARKETPLACES_SCHEMA TO ROLE AFTERSPOT_ETL_ROLE;
GRANT ALL ON ALL VIEWS IN SCHEMA MARKETPLACES_SCHEMA TO ROLE AFTERSPOT_ETL_ROLE;
GRANT ALL ON ALL TABLES IN SCHEMA RAW_SCHEMA TO ROLE AFTERSPOT_ETL_ROLE;
GRANT ALL ON ALL VIEWS IN SCHEMA RAW_SCHEMA TO ROLE AFTERSPOT_ETL_ROLE;

GRANT ALL ON FUTURE TABLES IN SCHEMA MARKETPLACES_SCHEMA TO ROLE AFTERSPOT_ETL_ROLE;
GRANT ALL ON FUTURE VIEWS IN SCHEMA MARKETPLACES_SCHEMA TO ROLE AFTERSPOT_ETL_ROLE;
GRANT ALL ON FUTURE TABLES IN SCHEMA RAW_SCHEMA TO ROLE AFTERSPOT_ETL_ROLE;
GRANT ALL ON FUTURE VIEWS IN SCHEMA RAW_SCHEMA TO ROLE AFTERSPOT_ETL_ROLE;

GRANT MONITOR, USAGE, CREATE TABLE, CREATE VIEW ON ALL SCHEMAS IN DATABASE AFTERSPOT_DB TO AFTERSPOT_ETL_ROLE;
```

RESULT #2 PYTHON SCRIPT FOR ETL

Content_Viewership		
CV_id	<pi>	Integer <M>
DeviceUser_id		Variable characters (50) <M>
UalpHash		Variable characters (100)
Content_id		Variable characters (50) <M>
Content_title		Variable characters (200)
Content_start		Date & Time
Content_end		Date & Time
Content_duration		Integer
Channel_id		Variable characters (50) <M>
Channel_name		Variable characters (100)
Session_Start		Date & Time
Session_End		Date & Time
Session_Duration		Integer
AdId		Variable characters (50)
BcId		Variable characters (50)
BcCategory		Variable characters (50)
CV_id	<pi>	

RESULT#3 VIEWERSHIP RELATION

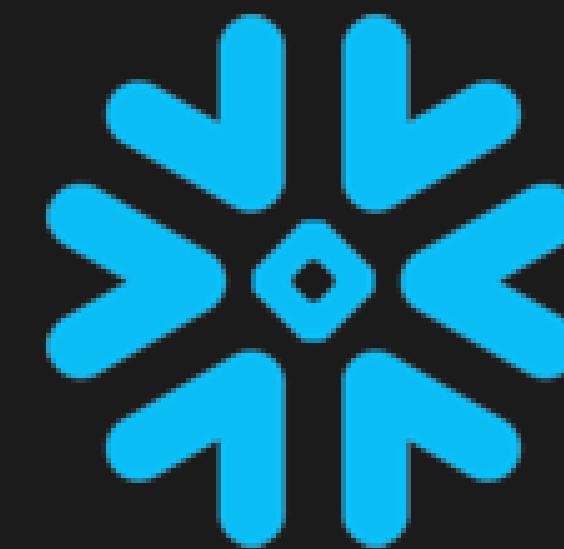
BUSINESS INTELLIGENCE

SERVICES

TECHNOLOGY
- II



SNOWFLAKE
DATA WAREHOUSE



DATA MARKET
PLACE

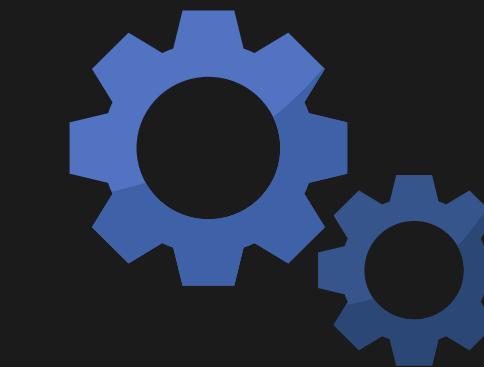
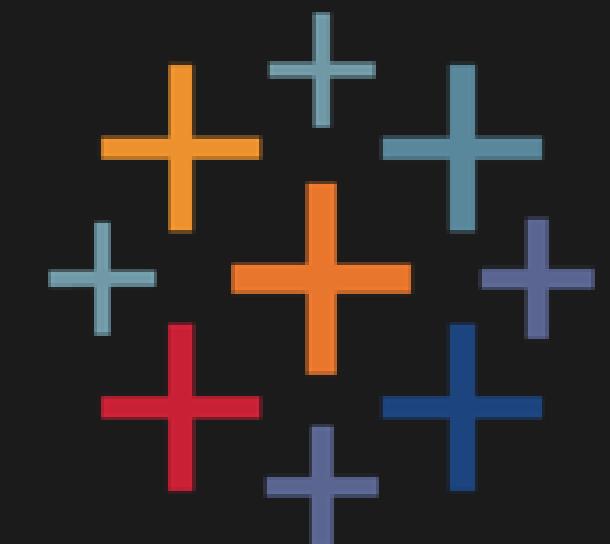


TABLEAU
BI VISUALIZATION



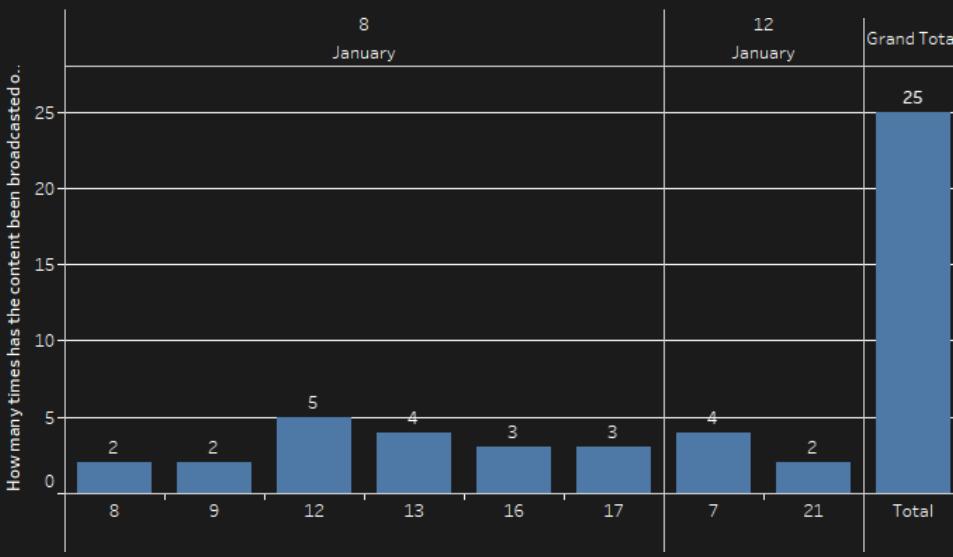
Create visualizations for different scenarios based on user behavior data



BUSINESS INTELLIGENCE

QUESTIONS - SCENARIOS

S2-c: Advertising spot "ČSOB"
Detect how many times the Ad has run on TV broadcast

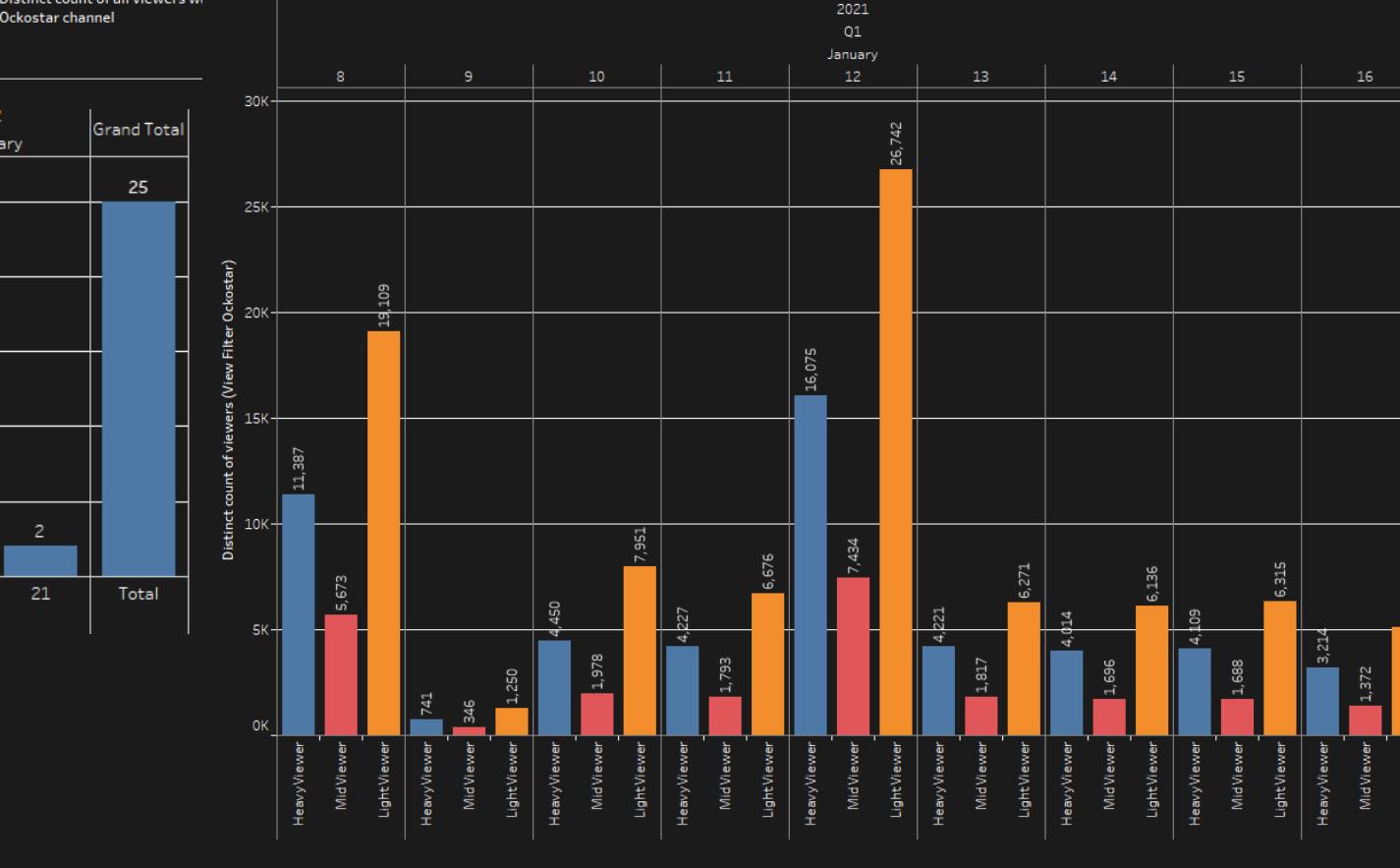


ANSWERS - VISUALIZATIONS

S1: Advertising spot "FIO BANKA"
Detect overall number of TV devices that - a) watched the Ad at least once b) did not watch the Ad even once (zero times)

Session Start (View Filter Ockostar)	2021										Grand Total
	8	9	10	11	January	12	13	14	15	16	
Distinct Viewers of Fiobanka Ad	7,521	622	2,802	2,407	7,516	1,249	2,518	1,163	952	12,847	

S6-a: TV devices activity/behaviour throughout the week
Detect all overall number of TV devices and categorize them based on how long watch TV channel.



RESULTS - II

Deliverables:

- Data marketplace set-up
- Answering questions in different scenarios in a real-life-like case study, i.e.:
 - Detecting the overall number of TV devices that - a) watched the Ad at least once b) did not watch the Ad.
 - Detecting all overall number of TV devices and categorizing viewers based on how long they are watching the TV channel.



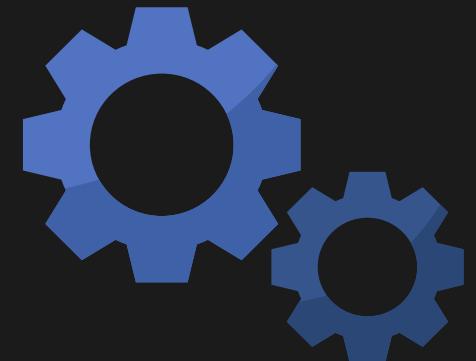
MACHINE LEARNING

WHAT ARE WE USING

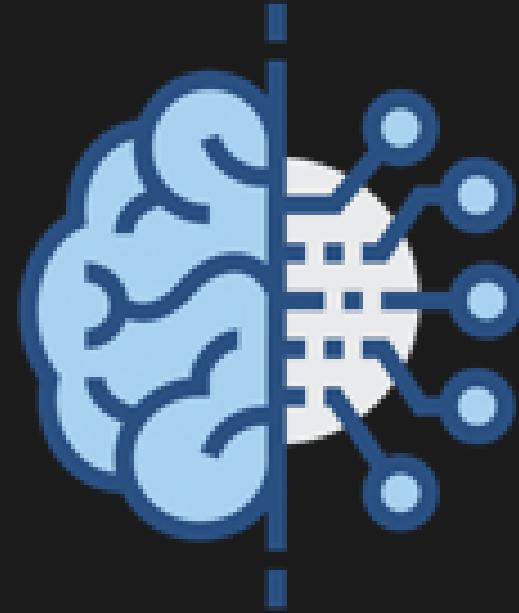
SNOWFLAKE
DATA WAREHOUSE



SAMPLE DATA



MACHINE LEARNING



TECHNOLOGY
- III

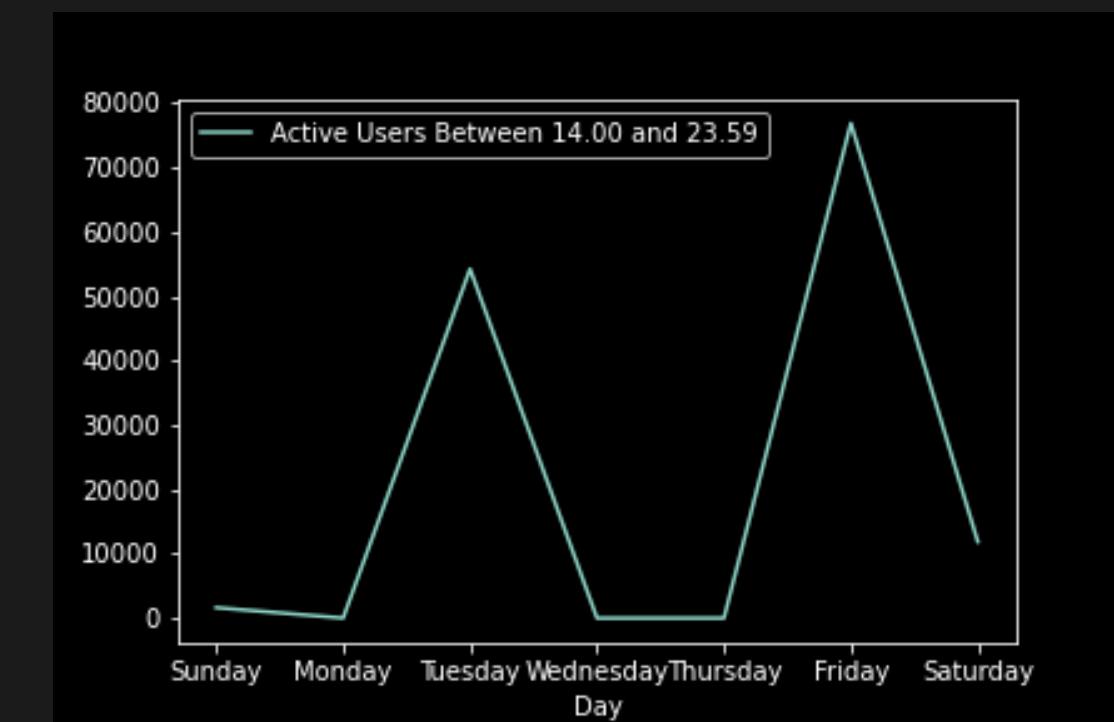
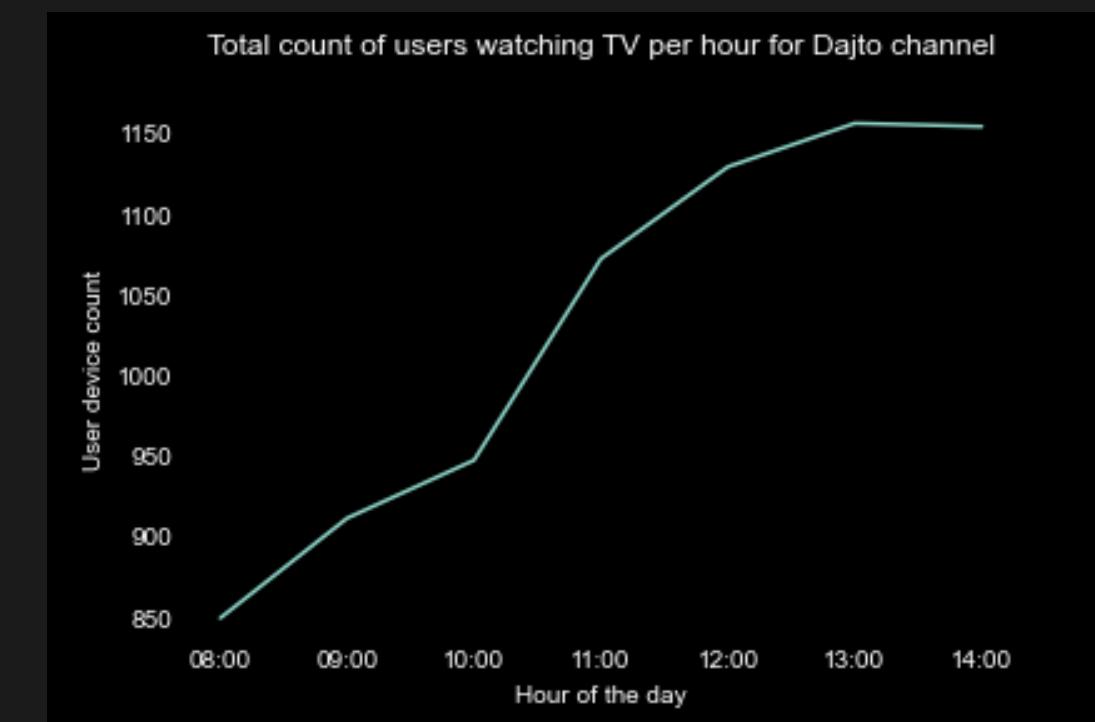
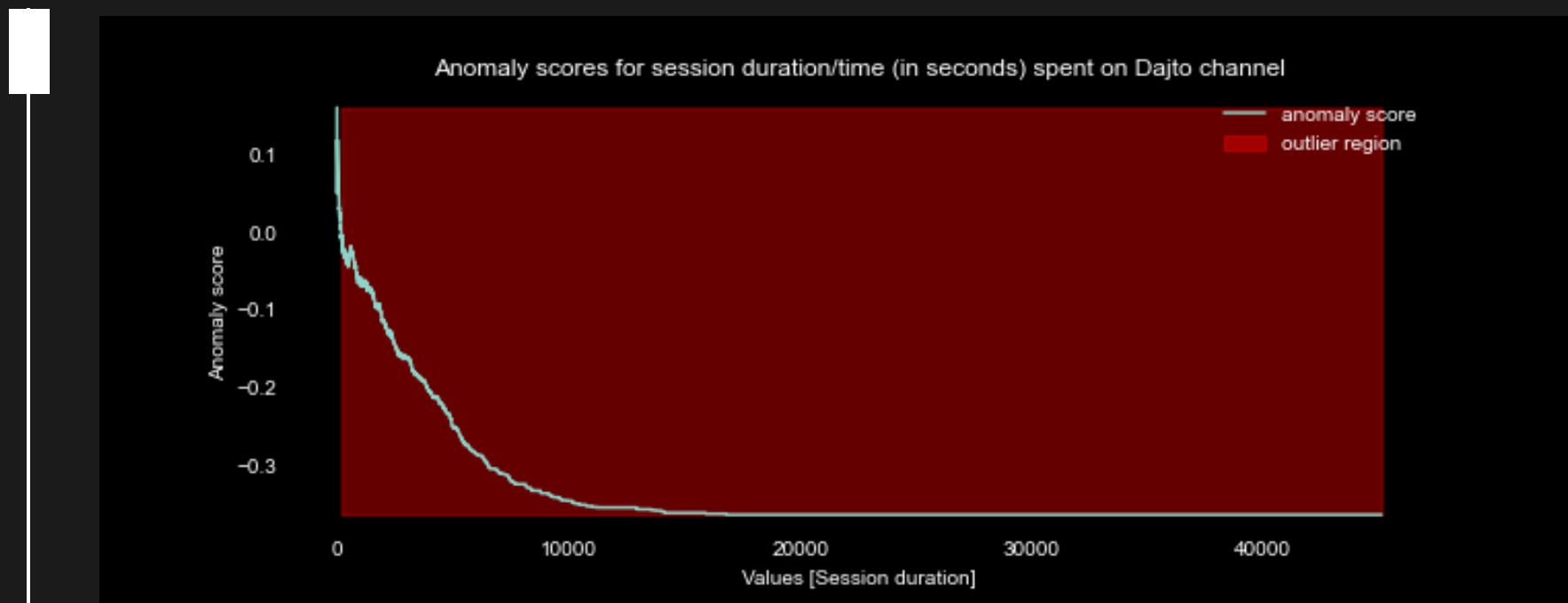
We programmed the computer to forecast the future of viewers' behavior



MACHINE LEARNING

Deliverables:

- 1) User traffic
 - Detect overall numbers and its time spend on the channel (traffic) and evaluate it on time period.
- 2) Create algorithm to help predict the amount of traffic (TV viewership) on time period.
- 3) Use an anomaly detection algorithm to predict and spot unusual user behavior per channel.



RESULT -
III

RESULTS



ACCOMPLISHED

- 1
- 2
- 3
- 4
- 5

Data architecture

3 data models, 3 database connectors (ETL), and 5 data tables.

BI processes

6 case scenarios.

Real scenarios to be executed

7/5 scenarios finished.

Data validation

11/6 finished.

Anomaly detection and forecasting with ML algorithms

forecast: 70% done [2/3 tasks].

TO BE DONE NEXT

- 1
- 2
- 3
- 4

DeepLearning for forecasting TV behavior.

100% personalized content and targeted advertising.

Infrastructure scaling to global level.

Maintain Facebook & Google independence.

PROCESS USED

ASSIGNMENT

Students are assigned a task and a support team from the company based on their skills.

DEVELOPMENT

Students work independently on each task and troubleshoot together.

REVIEW AND ADJUSTMENT

The team meets each day to assess progress and share information.

VALIDATION

Once a task is approved, the student is assigned a new task.

//19

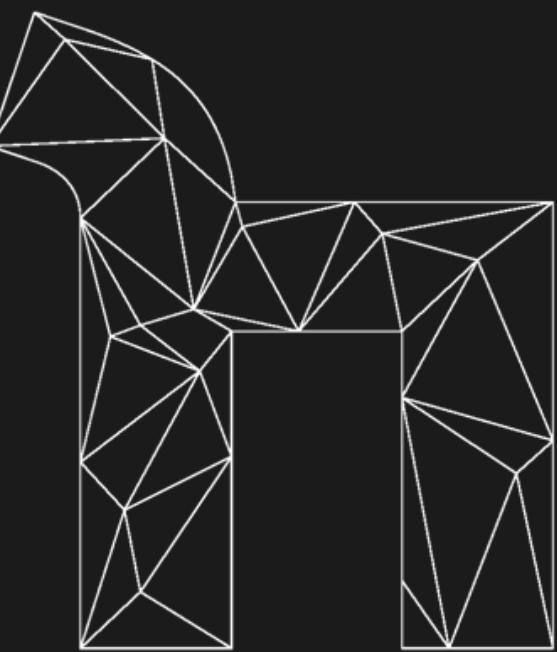
REFLECTIONS ON WHAT WE LEARNED

REAL | STRONG | FOCUSED | FUN | FOR ALL

- Application of theory in real-life project
- Always send updates no matter how little.
- Don't be afraid to ask questions.
- Resting isn't unproductive



//20



Q&A



THANK YOU!

DIGIHEALTH 008 - HYBRID