1. **Creating data sets and data cleaning on Data server**

* Command line (Bash) getting the data:

wget IPIPIPIPIPIPI/testtask/dilans\_data.csv

* Command Line (Bash) creating 3 data sets from the main: first readers, returning readers, subscribers, buyers

cat dilans\_data.csv | grep ’subscribe’ > /home/ownserver/dilans\_subs.csv

cat dilans\_data.csv | grep ’buy’ > /home/ownserver/dilans\_buy.csv

cat dilans\_data.csv | grep ’read’ > /home/ownserver/dilans\_read.csv

* Separating first readers and returning readers from all readers file:

grep -E 'Reddit|AdWords|SEO' dilans\_read.csv > dilans\_first\_read.csv

grep -E -v 'Reddit|AdWords|SEO' dilans\_read.csv > dilans\_more\_reads.csv

* Creating Date and Time variables from Timestamp:

cat dilans\_buy.csv | sed 's/ /;/' > dilans\_buy\_date.csv

cat dilans\_subs.csv | sed 's/ /;/' > dilans\_subs\_date.csv

cat dilans\_first\_read.csv | sed 's/ /;/' > dilans\_first\_read\_date.csv

cat dilans\_more\_reads.csv | sed 's/ /;/' > dilans\_more\_reads\_date.csv

cat dilans\_reads.csv | sed 's/ /;/' > dilans\_reads\_date.csv

1. **SQL Tables and Data load in Command Line**

psql -U XXXX -d postgres

CREATE TABLE dilans\_firstreader (

date DATE,

time TIME,

event\_type TEXT,

country TEXT,

user\_id BIGINT,

source TEXT,

region TEXT

);

CREATE TABLE dilans\_rereader (

date DATE,

time TIME,

event\_type TEXT,

country TEXT,

user\_id BIGINT,

region TEXT

);

CREATE TABLE dilans\_subs (

date DATE,

time TIME,

event\_type TEXT,

user\_id BIGINT

);

CREATE TABLE dilans\_purchase (

date DATE,

time TIME,

event\_type TEXT,

user\_id BIGINT,

amount INTEGER

);

\COPY dilans\_firstreader FROM '/home/ownserver/dilans\_first\_read\_date.csv' DELIMITER ';';

\COPY dilans\_rereader FROM '/home/ownserver/dilans\_more\_reads\_date.csv' DELIMITER ';';

\COPY dilans\_subs FROM '/home/ownserver/dilans\_subs\_date.csv' DELIMITER ';';

\COPY dilans\_purchase FROM '/home/ownserver/dilans\_buy\_date.csv' DELIMITER ';';

1. **Data analysis in PostgreSQL**

* Daily trends of users: 4 segments and within buyers creating the Ebook (spending 8) and Video (spending 80) subsegments
* Funnel analysis of users
* Daily KPIs overview
  + Daily active users
  + Daily revenue
* Composition of Ebook and Video buyer segments
  + Source, Country, Topic to read
  + Microsegments for communication fine-tuning
* ROI and Netnet income per buyer segment per source