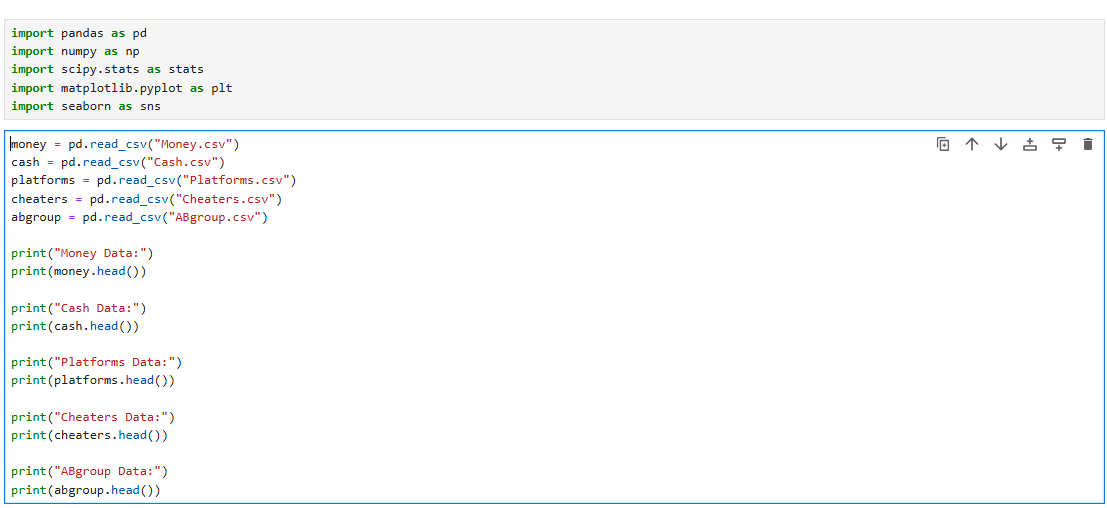
**Sirvin Iliasov  
Data analytics from scratch 2.0" course  
Final project  
  
Section 1: Purpose of the project  
What was the problem that needed to be solved?**Understand whether to run a discounted armor promotion in the future  
**How will you approach it?**  
Analyze data from five tables, clean up anomalies and cheaters, calculate metrics, compare ARPU (average revenue per player), **ARPPU (per paying player) and currency spend between test groups**, plot confidence intervals and visualize results.  
  
**Section 2: Analysis of sources**

Python was chosen( all code was written in Jupyter Notebook ) for several reasons:

1.I chose Python because it is the most convenient tool to analyze data using Pandas, NumPy.

2.Python allows you to quickly clean, merge, and visualize data.



**Section 3: Data cleansing**

We used a **dual method of filtering out cheaters**:

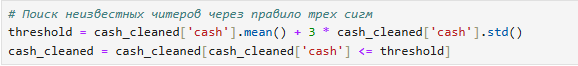
Removal of known cheaters (according to the list from cheaters.csv):

Excluded all users who have **flag = 1** in the cheaters table (i.e. they are definitely cheaters). 

**Search for unknown cheaters** (by abnormal spending of in-game currency):

We used **the three sigma rule** to identify anomalies.

If a player **spends significantly more** in-game currency than 99% of players, we delete them.

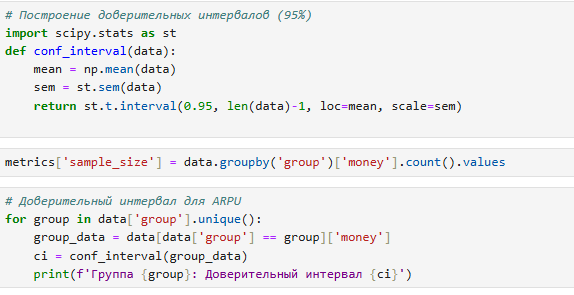


**Bottom line**: We've removed **all known cheaters** and **weeded out the abnormally suspicious** ones by spending game currency.

**Section 4: Using Statistical Methods**

The confidence interval indicates the range in which the true mean value can lie with a **95% probability.**

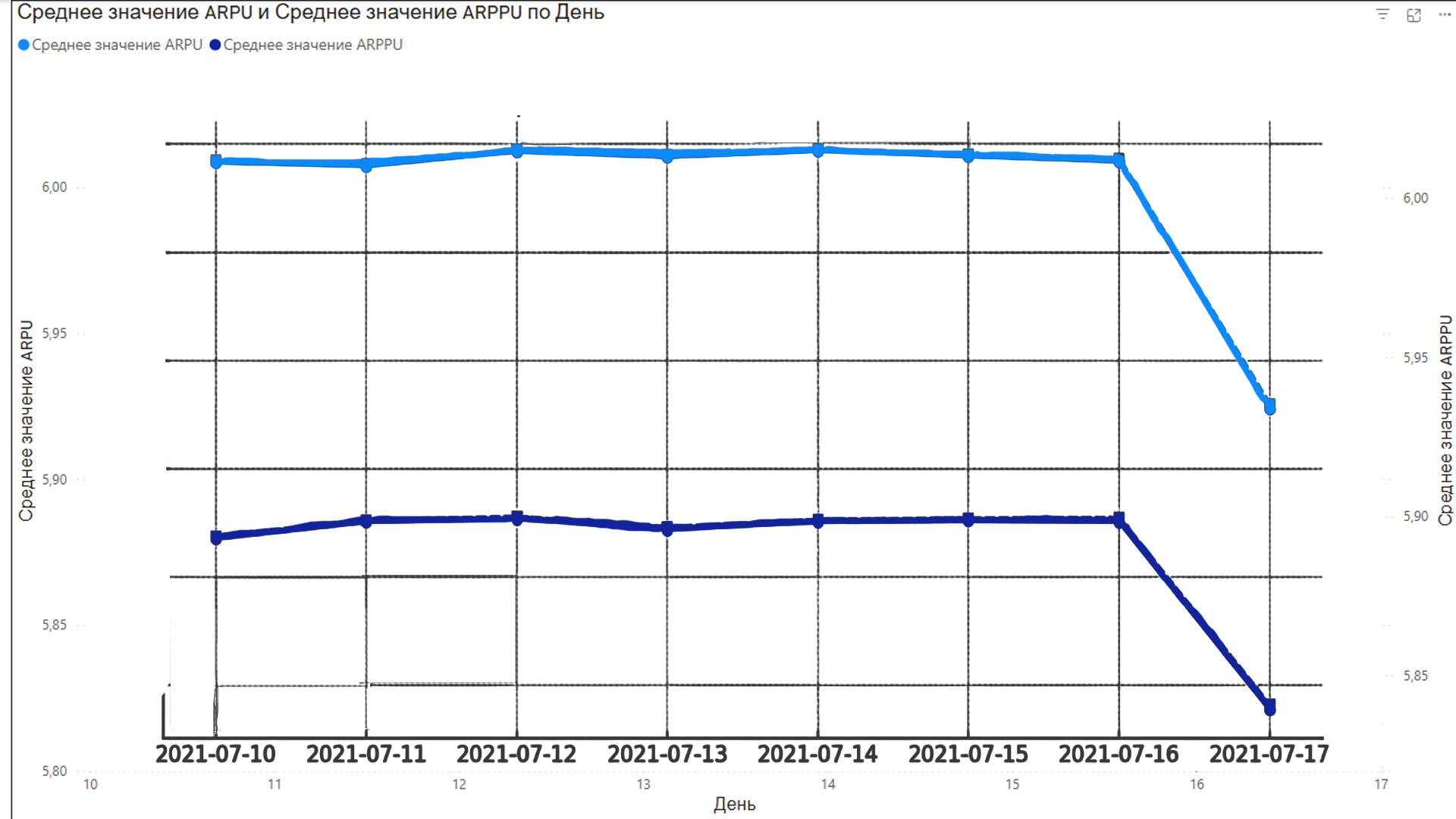
Code for calculating confidence intervals:



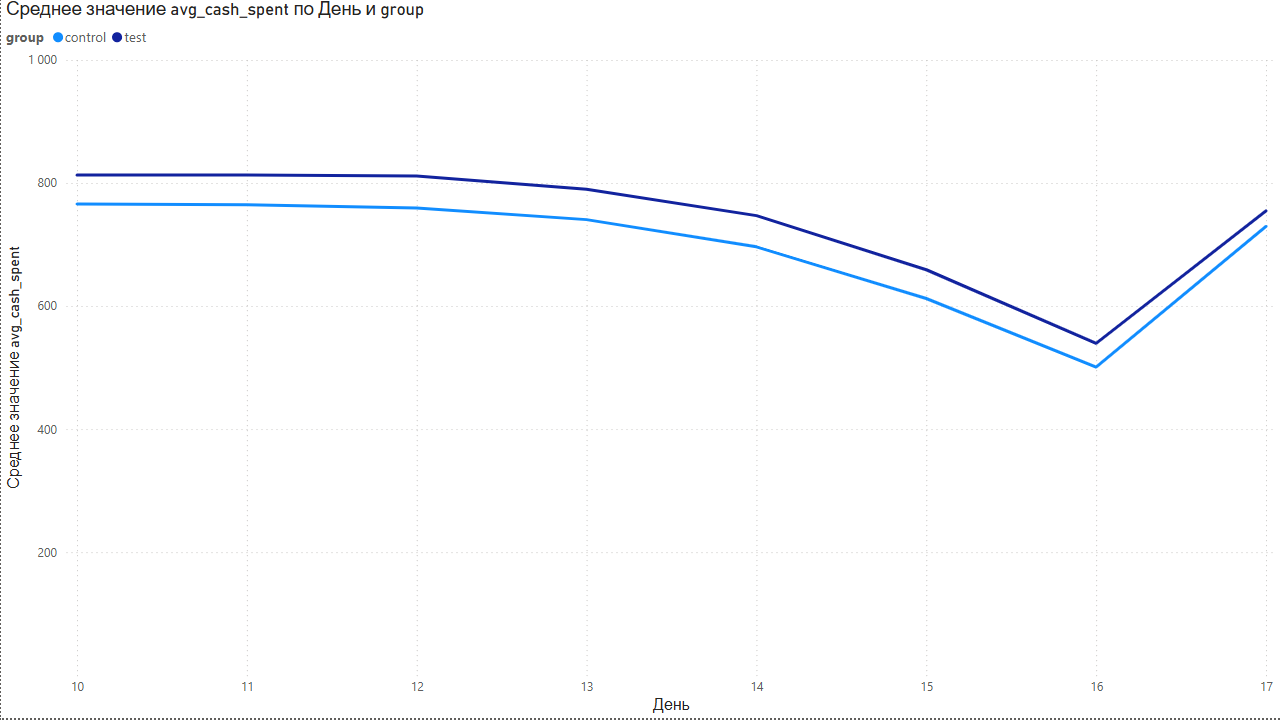
**Section 5: Report generation**Building graphs in Power BI

We will now plot **2 graphs.**

**1. Line graph of ARPU and ARPPU by day.This graph will show how ARPU and ARPPU changed in the control and test groups.**

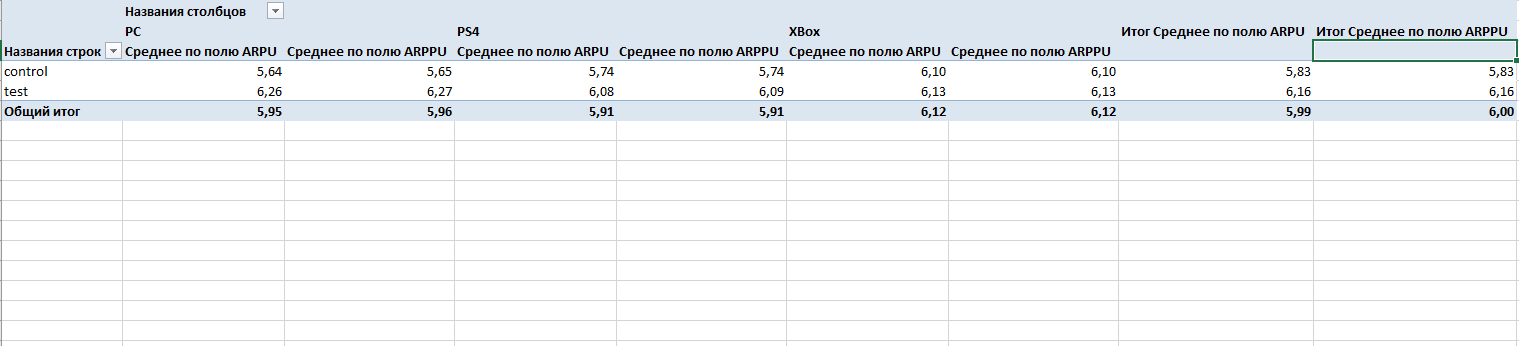


**2: A line graph of in-game currency spending by day.** This graph will show **how in-game money (Cash) was spent in the test and control groups.**



**avg\_cash\_spent** is the **average spending of in-game currency** per day per player

Excel summary table with ARPU by group and platform.



Bottom line

The campaign can be repeated, but with improvements. If the goal is **maximum profit growth**, it is worth **optimizing the mechanics of the promotion** before re-launching it, because as long as the difference in metrics is **small,** the effect may be random.