

David's Travel Guide analysis

by Andras Petre

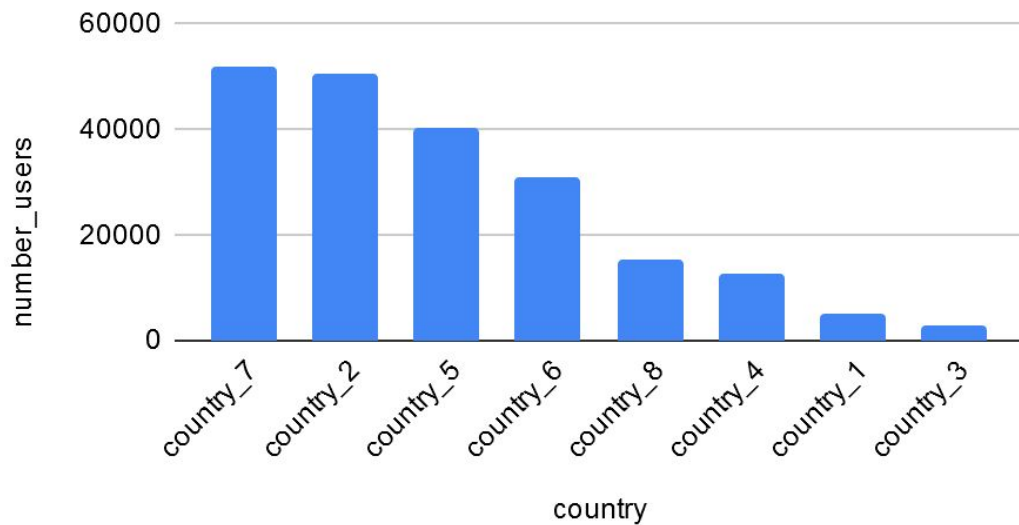
A large, dark blue, diagonal shape that starts from the bottom left and extends towards the top right, covering the lower half of the page.

Introduction

- David writes a blog about his travel experiences. He writes about different topics, his blog is available in 8 countries. He finance his blog from different sources in different rate
- The dataset contains over 600 000 rows
- This project aims is to segment the data and discover some statistical values in the dataset and give advice David e.g. in which country he should prioritise his effort.
- The main findings in the data have been presented in graphical format and the data analysis has been carried out in Python and SQL
- The Python libraries used include Pandas, Numpy and Matplotlib

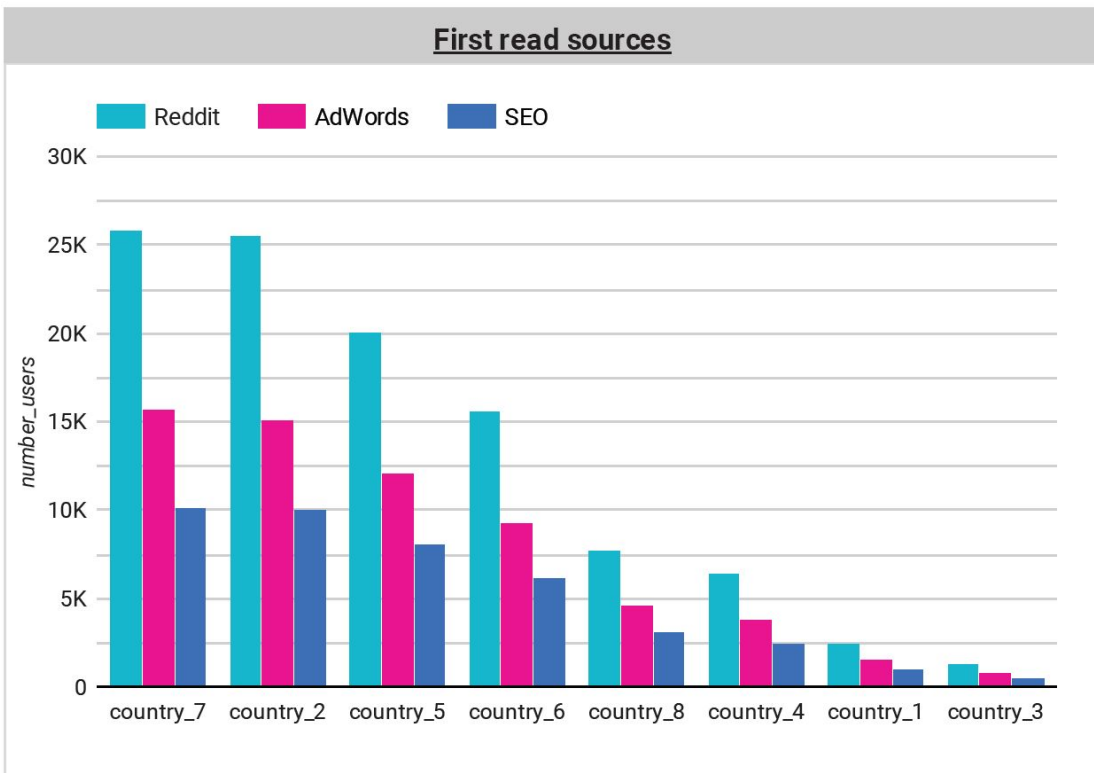
First read distribution 1.

first read users by country



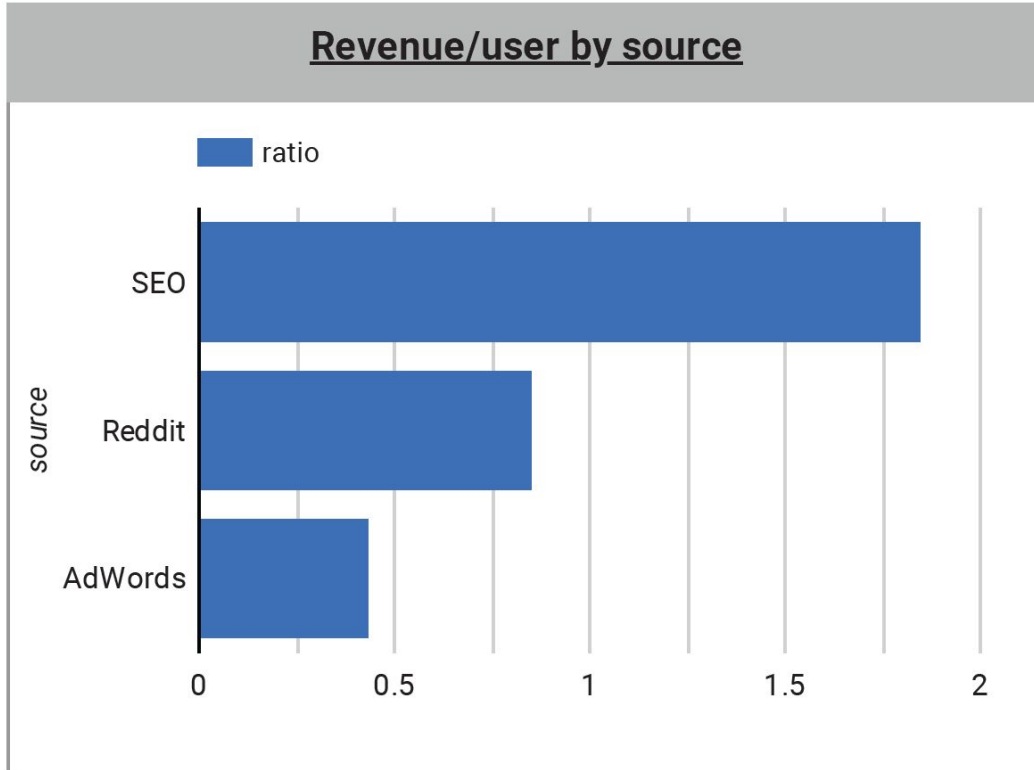
- The most first readers come from country_7, and country_2, so he should focus on them
- Country_1 and country_3 have the less new readers

First read distribution 2.



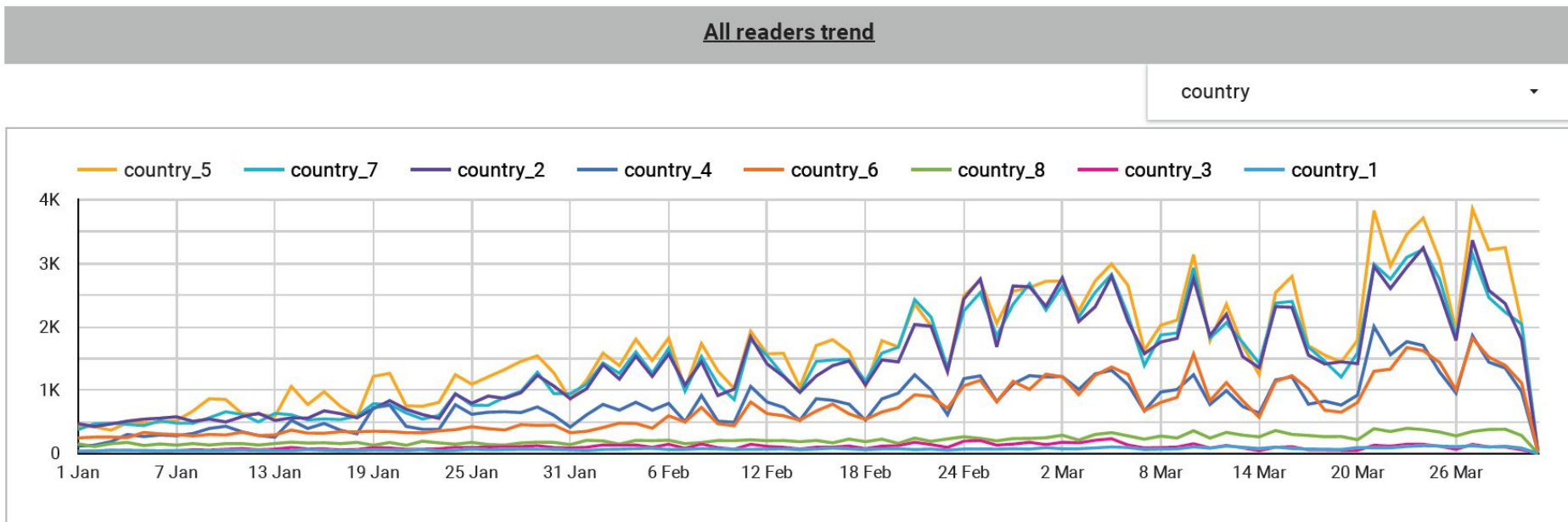
- By number of new readers Reddit the top source, so he should more invest into Reddit to gain more new visitors on his page
- ...and at the same time David should have to consider to reduce the investment into SEO

Revenue/user ratio



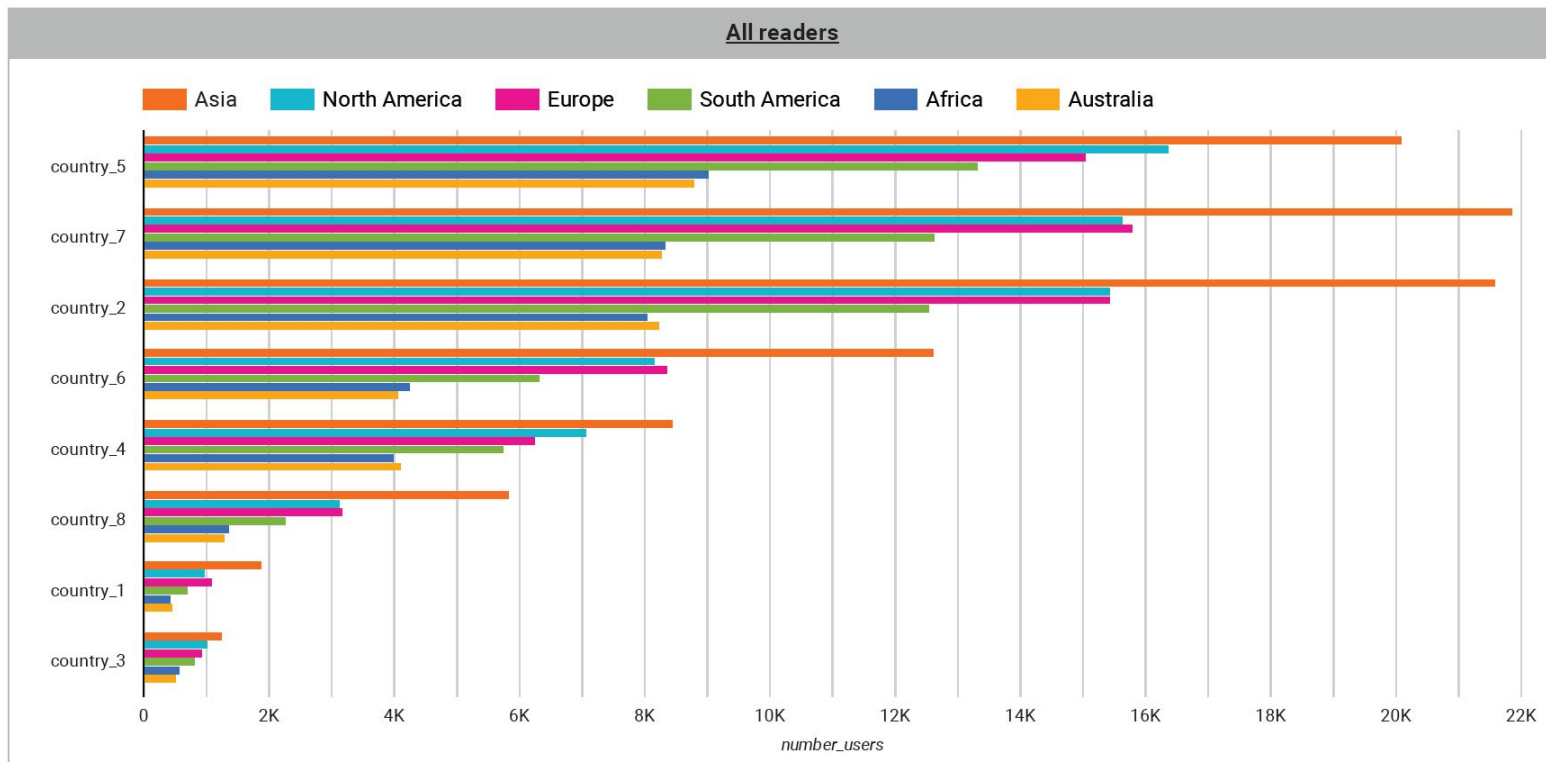
- In the conversation rate is SEO the most valuable source for his travel blog
- If it's more important for him, than the number of new readers, he should invest more into SEO, and less into AdWords

All readers trends



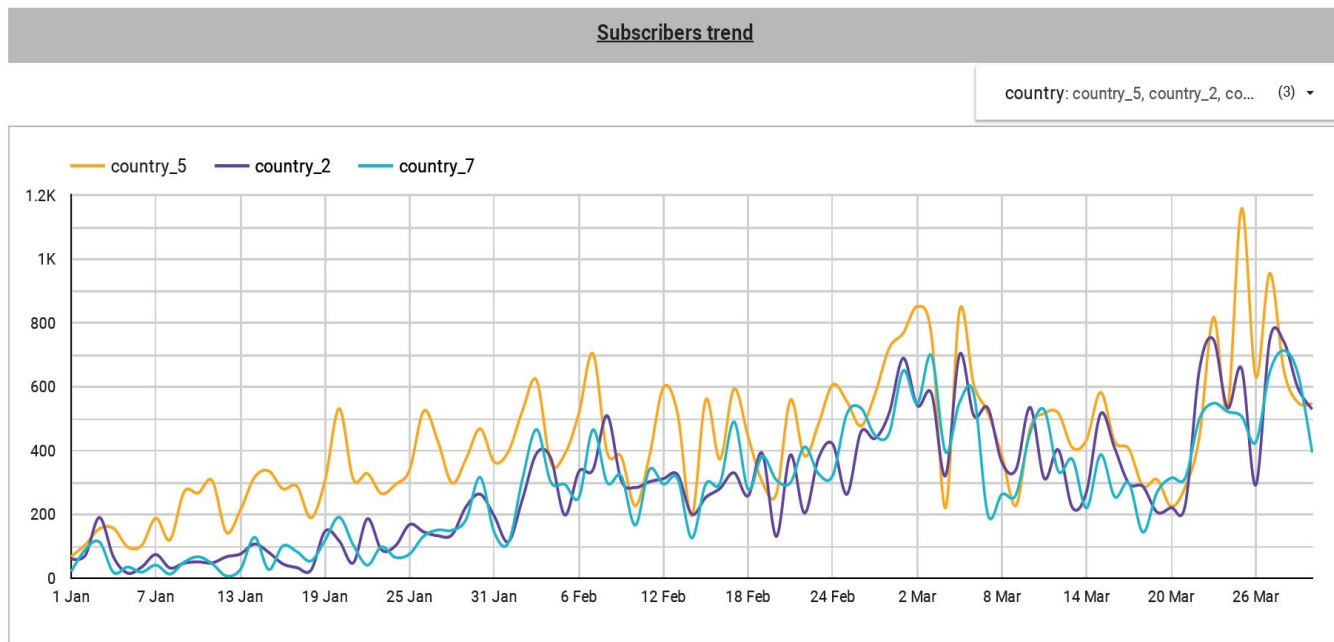
There are small groups (2-3) of countries which trends “move” together in some periods e.g. (country_5, 7, 2), so he should focus on this countries

Readers distribution by topic



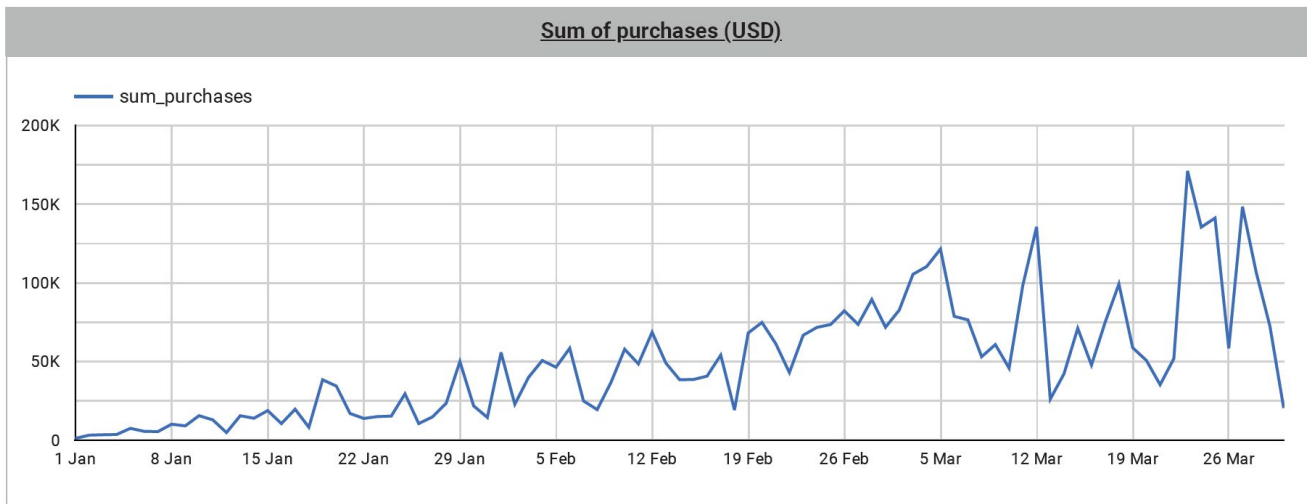
The most popular continents are Asia by all countries, so he should create more articles and videos about countries in Asia to gain more readers in the future

TOP 3 country subscribers trend



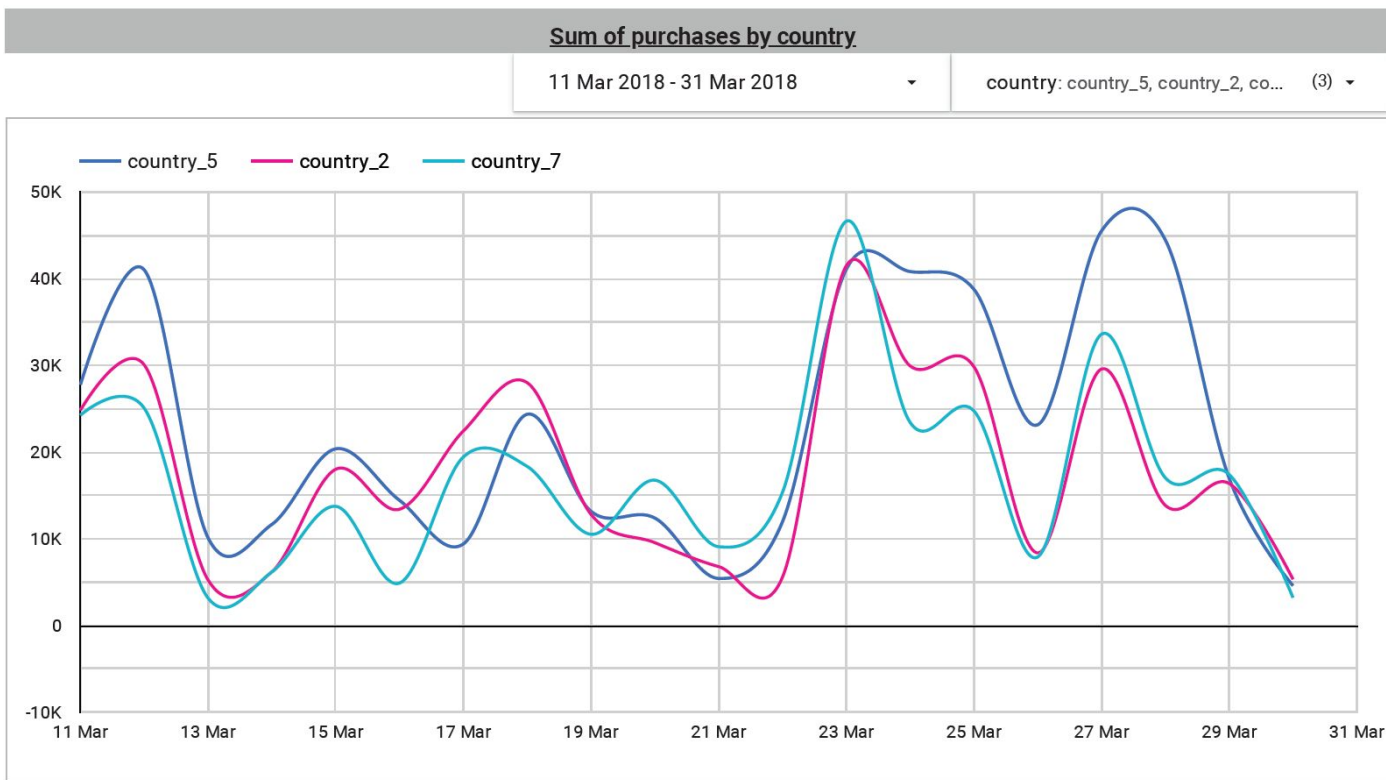
- The most subscribers come from the same countries as by the readers, so by them are David's blog very popular
- We can see a slightly growth over the time, however there are suddenly crashes in some periods

All purchases trend



We can see the similar trends in the purchases as in the previous charts according to the readers and subscribers: slightly increase on the whole and some sudden crashes

TOP 3 country purchases trend (last 20 days)



- We can see a quite waving chart by the top 3 countries
- The top 3 countries are the same as by the readers and by the subscribers (5, 2, 7)

Ratio of readers, subscribers, purchasers 1.

Only a little part of the readers subscribe and buy a product:

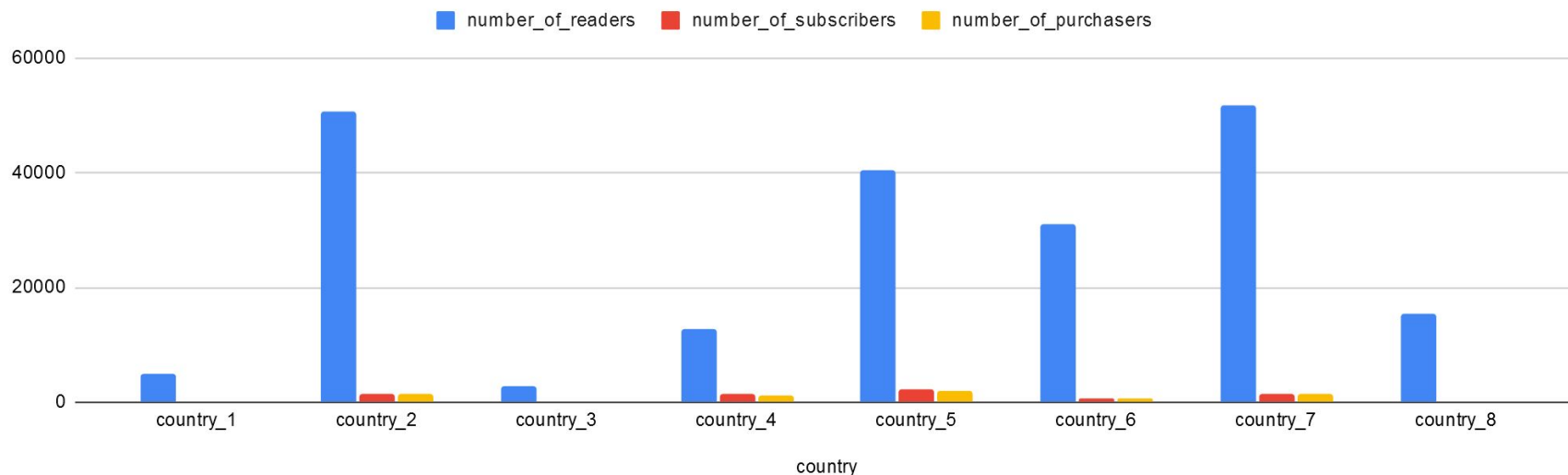
country	number_of_readers	number_of_subscribers	number_of_purchasers	subscribers/readers	purchasers/readers
country_1	5076	28	25	0.55%	0.49%
country_2	50675	1558	1443	3.07%	2.85%
country_3	2742	160	112	5.84%	4.08%
country_4	12751	1419	1096	11.13%	8.60%
country_5	40349	2334	1971	5.78%	4.88%
country_6	31156	546	535	1.75%	1.72%
country_7	51791	1470	1381	2.84%	2.67%
country_8	15483	103	85	0.67%	0.55%

- The most readers come from country_7
- The most subscribers come from country_5, but by conversation is country_4 is on the top
- The most purchasers come from country_5 as well

Ratio of readers, subscribers, purchasers 2.

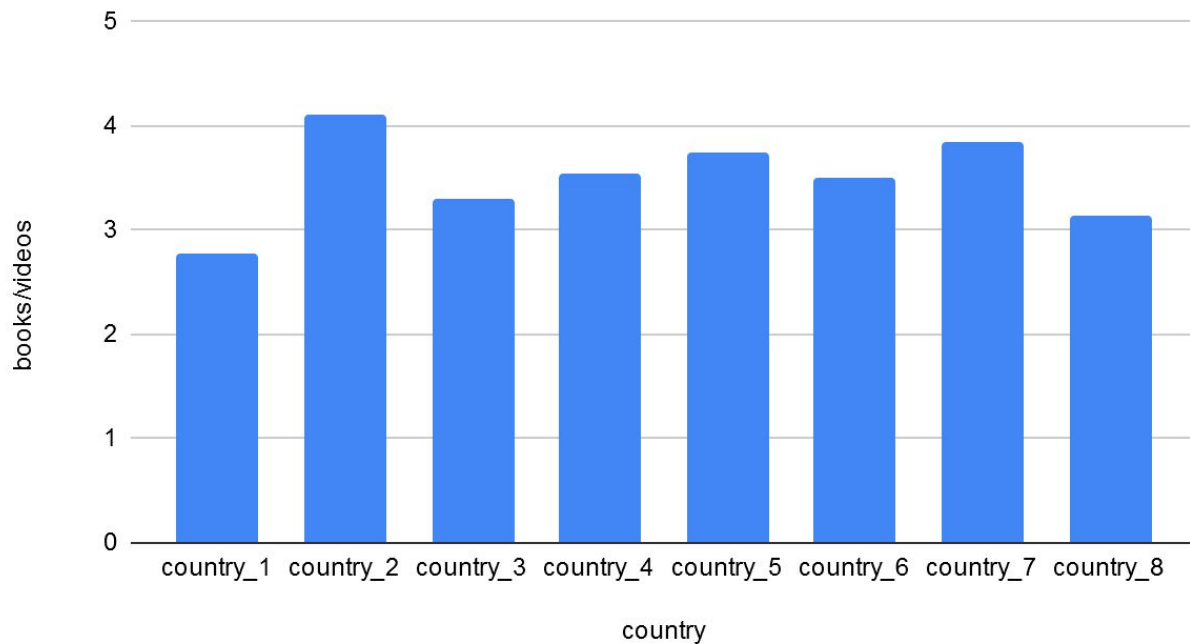
The previous table in a bar chart:

number_of_readers, number_of_subscribers and number_of_purchases



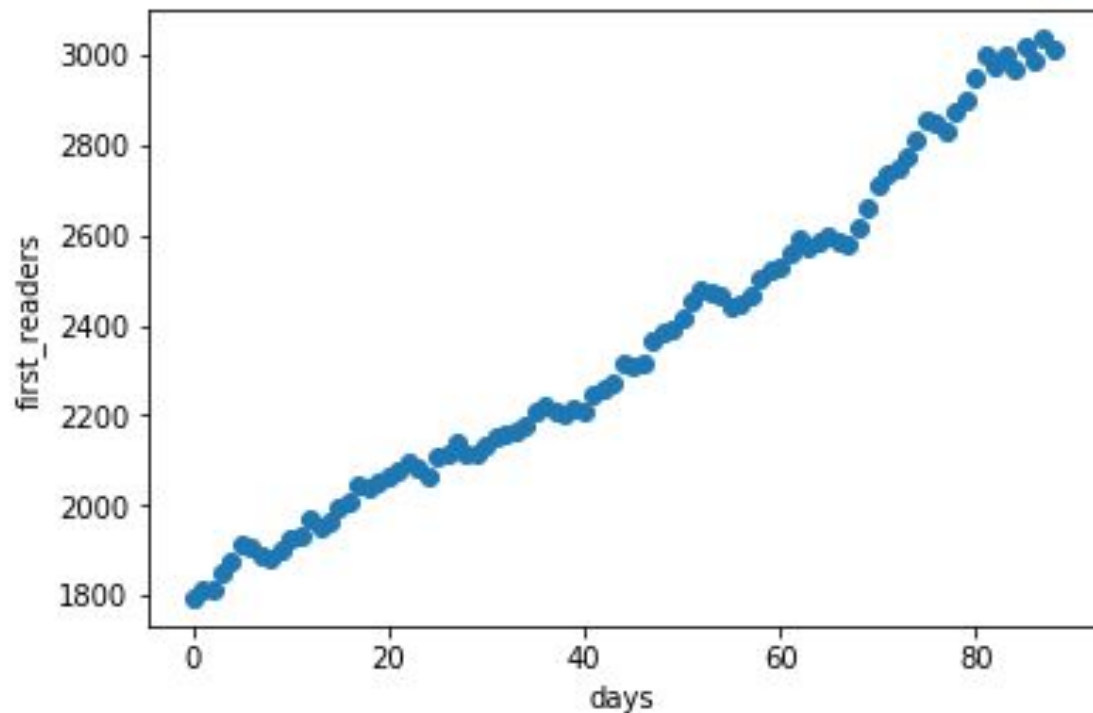
Ratio products by country

Books/videos ratio



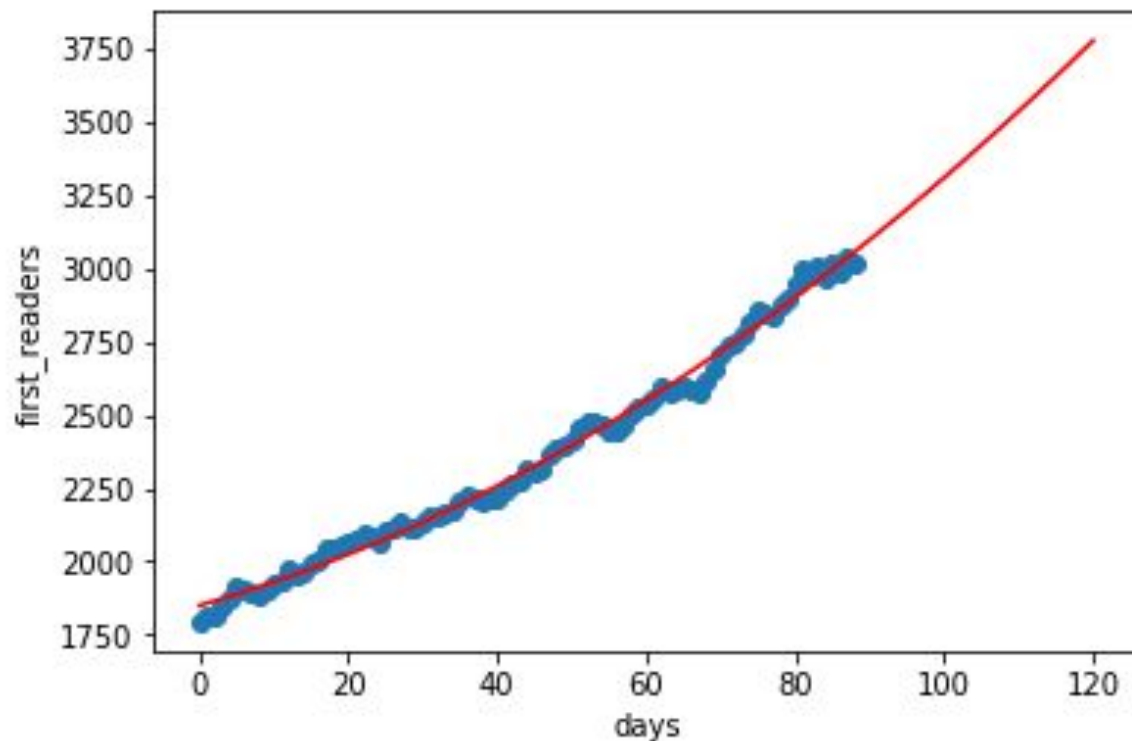
Quite similar ratio (3-4 times more books) between books and videos in almost all countries, exception is country_1 and country_2

Prediction (linear regression) 1.



- David's blog has about 3000 readers in the first 3 months
- His goal is to reach the 4000 first readers after 120 days
- Is it possible?

Prediction (linear regression) 2.



According to this prediction he can't reach his goal, but after a couple of days he is going to have the 4000. first reader on his blog :)

Conclusion

- David should focus on three country (country_7, country_2, country_5) especially, if want to increase the number of his readers and the revenue as well.
- By number of new readers Reddit the top source, so he should more invest into Reddit to gain more new visitors on his page. On the other side by the conversation rate is SEO the most valuable source for his travel blog.
- The most popular continents are Asia by all countries, so he should create more articles and videos about countries in Asia to gain more readers and subscribers in the future

Thank you

Contact details: andras.petre@gmail.com

A dark blue diagonal gradient bar that starts from the bottom left corner and extends towards the top right corner, covering the lower half of the slide.