Author: Ing. Denis Kopásek

DEMONSTRATION

INTRO

This document showcases my abilities in data analysis and concise presentation. It highlights two key projects:

- Web Traffic Analysis: This section delves into the analysis of traffic data, including web address collection and visitor behavior patterns focusing on repeat visitors.

- Material Engineering Experimentation: This section summarizes the processed data from experiments conducted in the field of material engineering, along with a comparative analysis.
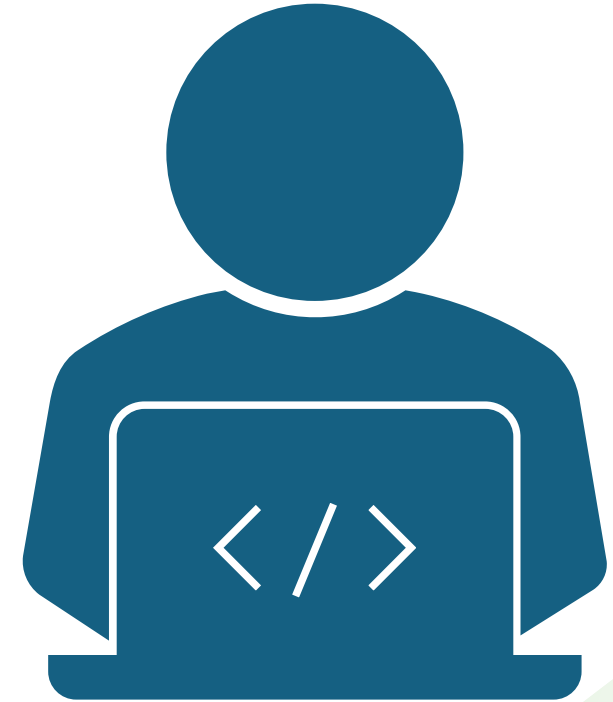
TASKS

QUESTIONS (PART A)

1. What percent of customers returned after the first visit?

2. What are the top three best-performing stories in each section, by pageviews?

3. Based on this data, would you choose to promote a Tech story or a Markets story on social media? Why?

4. Create a visualization exploring the relationship between any of the content characteristics (such as section, author, keywords etc...) and returning visitors. This is an open-ended task. Briefly describe the visualization and the insight.

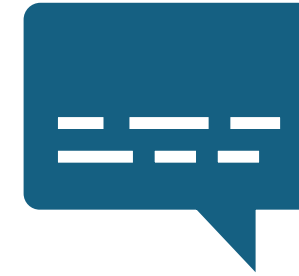5. What other interesting stories can you tell with this data?

SOLUTIONS
Part A

Q1

Question #1: What percent of customers returned after the first visit?

Answer #1: Percentage of customers who returned after the first visit: 12.8%

Q2

**Question #2:** What are the top three best-performing stories in each section, by pageviews?

Answer #2:

Section	Headline	Pageviews
WSJ_Business	How Your Returns Are Used Against You at Best Buy, Other Retailers	455
WSJ_Business	Toys 'R' Us Tells Workers It Will Likely Close All U.S. Stores	179
WSJ_Business	Are You Underpaid? In a First, U.S. Firms Reveal How Much They Pay Workers	130
WSJ_Life	The Truth About the SAT and ACT	121
WSJ_Life	How Kubrick's '2001: A Space Odyssey' Saw Into the Future	121
WSJ_Life	The Hottest Social Scene in Town Isn't the Singles' Bar. It's the Sup	100
WSJ_Markets	Intel Considers Possible Bid for Broadcom	137
WSJ_Markets	Do You Own Bitcoin? The IRS Is Coming for You	114
WSJ_Markets	Lloyd Blankfein Prepares to Exit Goldman Sachs as Soon as Year's End	71
WSJ_Opinion	The Exhaustion of American Liberalism	137
WSJ_Opinion	The Trump Tariff Layoffs Begin	74
WSJ_Opinion	Andrew McCabe and Consequences	57
WSJ_Politics	Trump's Personal Assistant Is Fired	1007
WSJ_Politics	Trump Lawyers Seek Deal With Mueller to Speed End of Russia Probe	664
WSJ_Politics	Andrew McCabe Kept Notes About Conversations With Trump, Gave Them t	159
WSJ_Tech	Facebook Really Is Spying on You, Just Not Through You	102
WSJ_Tech	Why Blockchain Will Survive, Even If Bitcoin Doesn't	49
WSJ_Tech	Why Intel Is So Wary of a Broadcom-Qualcomm Merger	40

Q3 (1/2)

Question #3: Based on this data, would you choose to promote a Tech story or a Markets story on social media? Why?
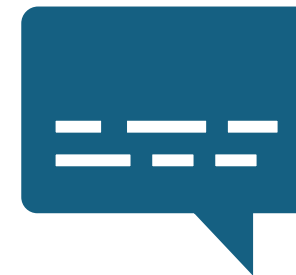
Answer #3.1: I would choose to promote a Tech story on social media. The Tech section already generates significant interest, with frequent mentions of social media-related keywords like Facebook, Twitter, and Instagram. This suggests strong existing engagement.

In contrast, the Markets section has fewer mentions of social media, indicating less engagement. Promoting Tech stories is likely to yield better results. However, if the goal of our company is to diversify content across sections, promoting a Markets story could be beneficial.

Nonetheless, my primary recommendation is to focus on the Tech section.

Q3 (2/2)

Question #3: Based on this data, would you choose to promote a Tech story or a Markets story on social media? Why?
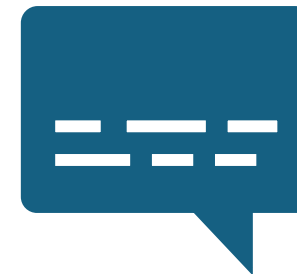
Answer #3.2: Speaking of numbers, analyzing the performance metrics of each category, we observe that the Tech section recorded a total of 704 individual visits, of which 74 resulted in subsequent visits, representing a return rate of 10.5%.
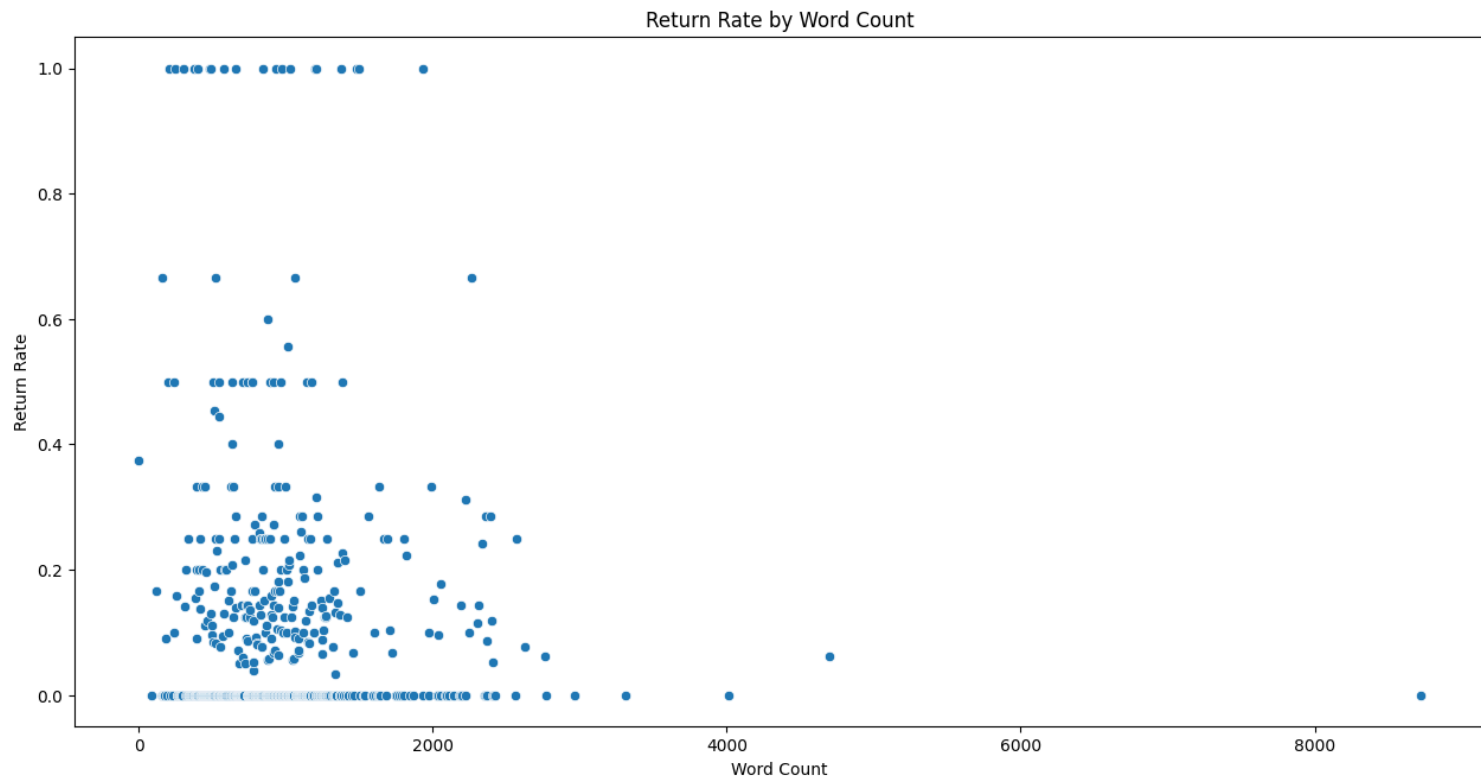
Conversely, the Markets section saw 992 unique visits, with 118 of these visits being repeat visits, corresponding to a return rate of 11.9%.

Based on these figures, it is evident that the **Markets section** not only **attracts a larger audience** but also **has a higher retention rate**. Consequently, it would be strategically advantageous to promote content from the Markets section, given its broader reach and greater likelihood of engaging returning visitors.

Q4 (1/2)

Question #4: Create a visualization exploring the relationship between any of the content characteristics (such as section, author, keywords etc...) and returning visitors. This is an open-ended task. Briefly describe the visualization and the insight.
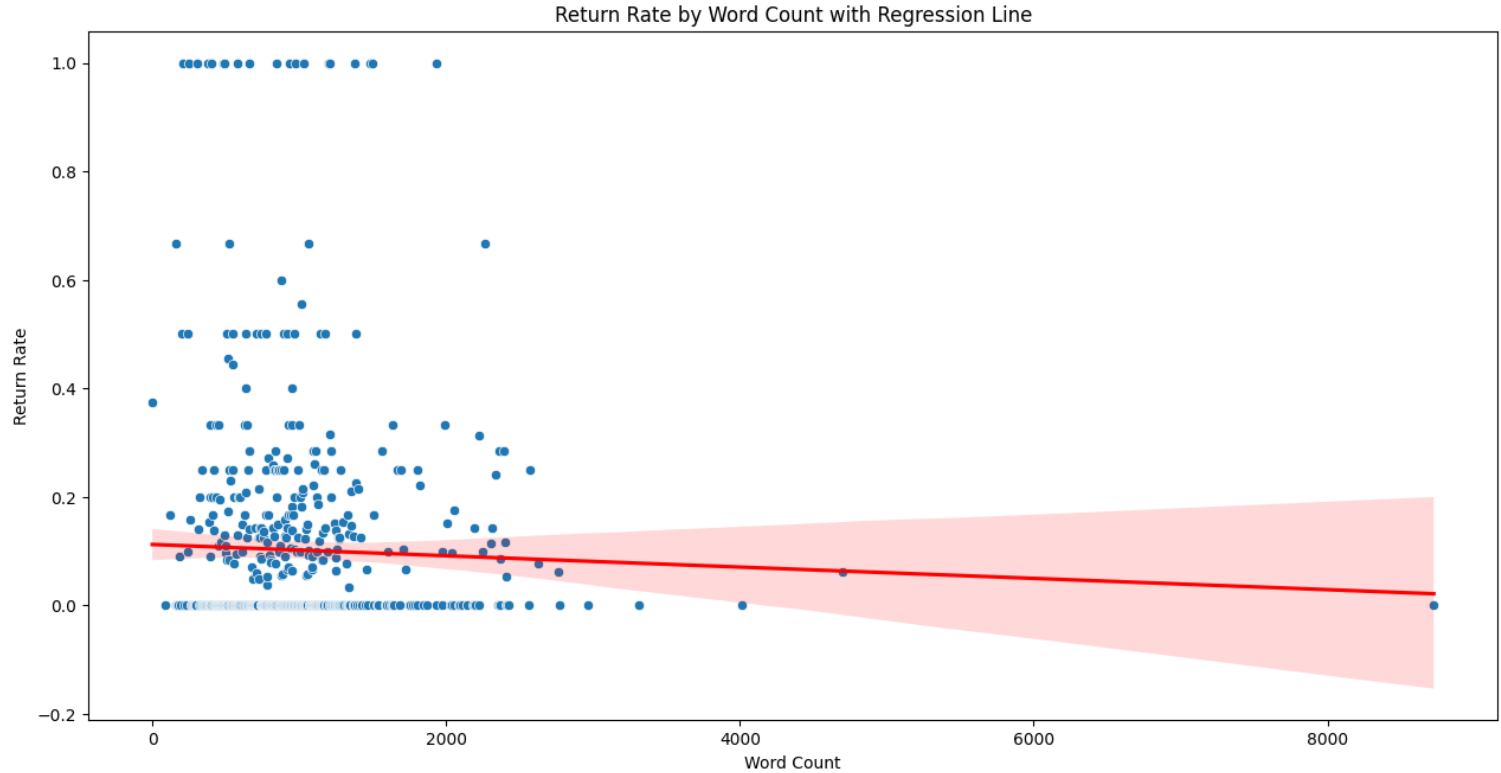

Return Rate by Word Count

Answer #4.1: After analyzing the relationship between word count and return rate, the correlation coefficient was found to be close to 0, indicating a **weak or no linear relationship**. Additionally, the scatter plot with a regression line showed a dispersed pattern of points around the line, further supporting the lack of a clear correlation. Therefore, it can be concluded that **word count does not** significantly **affect the return rate** of users.

Q4 (1/2)

Return Rate by Word Count with Regression Line

Question #4: Create a visualization exploring the relationship between any of the content characteristics (such as section, author, keywords etc...) and returning visitors. This is an open-ended task. Briefly describe the visualization and the insight.

Answer #4.1: After analyzing the relationship between word count and return rate, the correlation coefficient was found to be close to 0, indicating a **weak or no linear relationship**. Additionally, the scatter plot with a regression line showed a dispersed pattern of points around the line, further supporting the lack of a clear correlation. Therefore, it can be concluded that **word count does not** significantly **affect the return rate** of users.

Q4 (2/2)

Question #4: Create a visualization exploring the relationship between any of the content characteristics (such as section, author, keywords etc...) and returning visitors. This is an open-ended task. Briefly describe the visualization and the insight.

Return Rate by Section
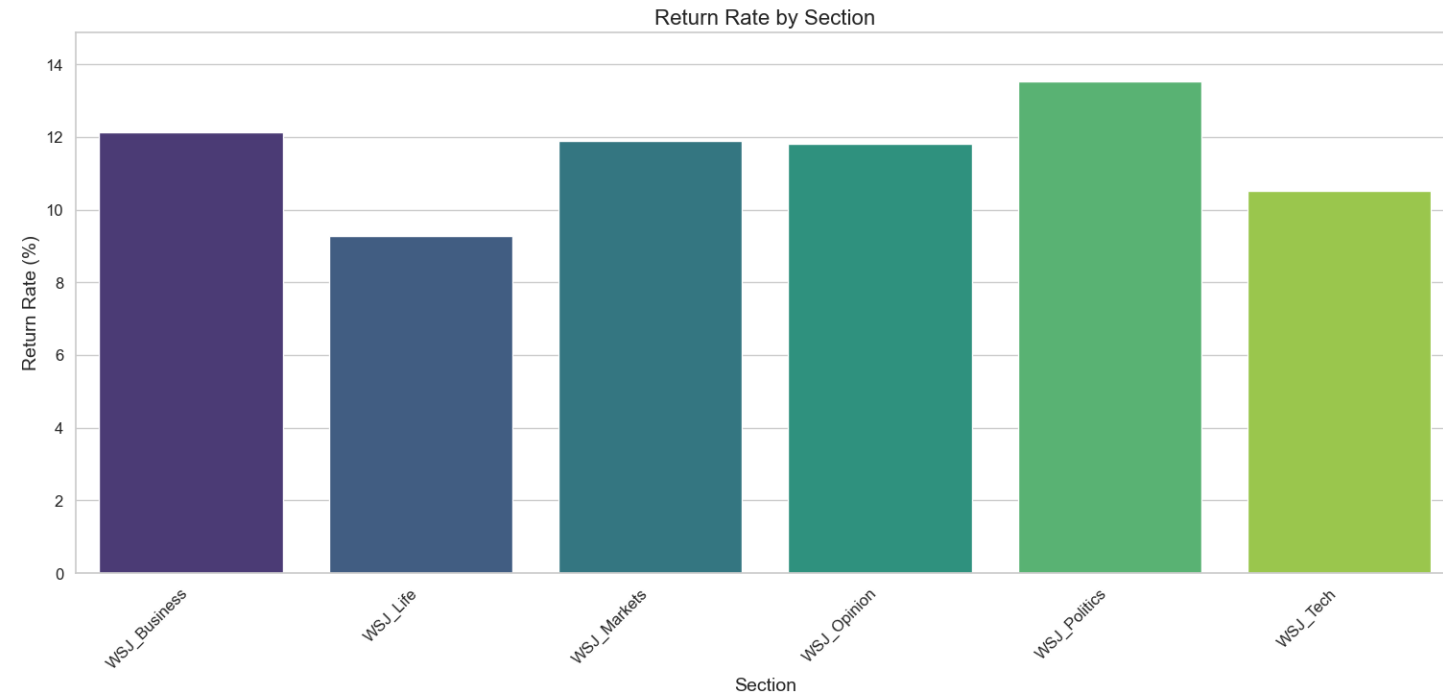
Answer #4.2: The analysis indicates that the **return rate** of visitors **does not** significantly **vary based on the section** they visited during their first visit. The return rates range from a **maximum of 13.5%** in the WSJ_Politics section to a **minimum of 9.3%** in the WSJ_Life section. Other sections have return rates that fall evenly within this range.

Q5 (1/4)

Question #5.1: What interesting stories can you tell with this data?

Word Cloud for WSJ_Business

Answer #5.1: Based on the top five performing articles in each section (based on amount of appearance in headline), we can identify overlapping keywords. By focusing on these keywords, we can strategically suggest and promote articles that attract on topics of articles in relation to those keywords to attract attention of first-time visitors.
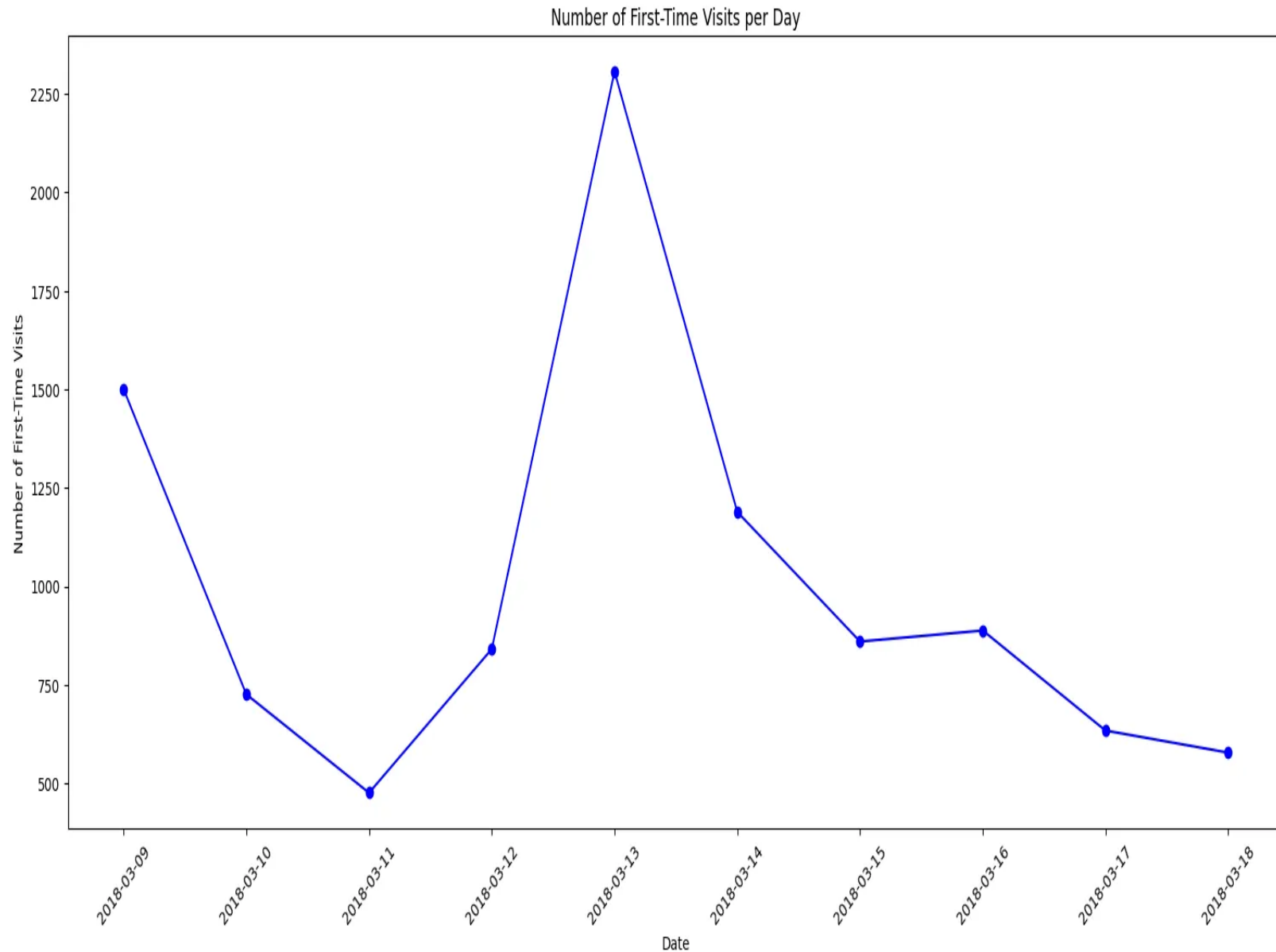
Q5 (2/4)

Question #5: What interesting stories can you tell with this data?

Answer #5.2: The graph indicates significant spikes in attention on March 9, 2018, and March 13, 2018. I recommend investigating the events or stories that drove this increased interest (keywords). Based on the results we can focus on future strategy and try to predict visitors' behavior in relation to the similar content in the future.

Number of First-Time Visits per Day

Q5 (3/4)

Question #5: What interesting stories can you tell with this data?

Answer #5.3: The graph indicates significant spikes in attention on March 9, 2018, and March 13, 2018. Upon further investigation of daily viewership by section, it appears that the Politics section was the primary driver of this increased traffic to the website. I recommend investigating the events of that day and focusing on similar stories to replicate this success.

Number of First-Time Visits by Section for Each Date (Stacked)

Section
- WSJ_Business
- WSJ_Life
- WSJ_Markets
- WSJ_Opinion
- WSJ_Politics
- WSJ_Tech
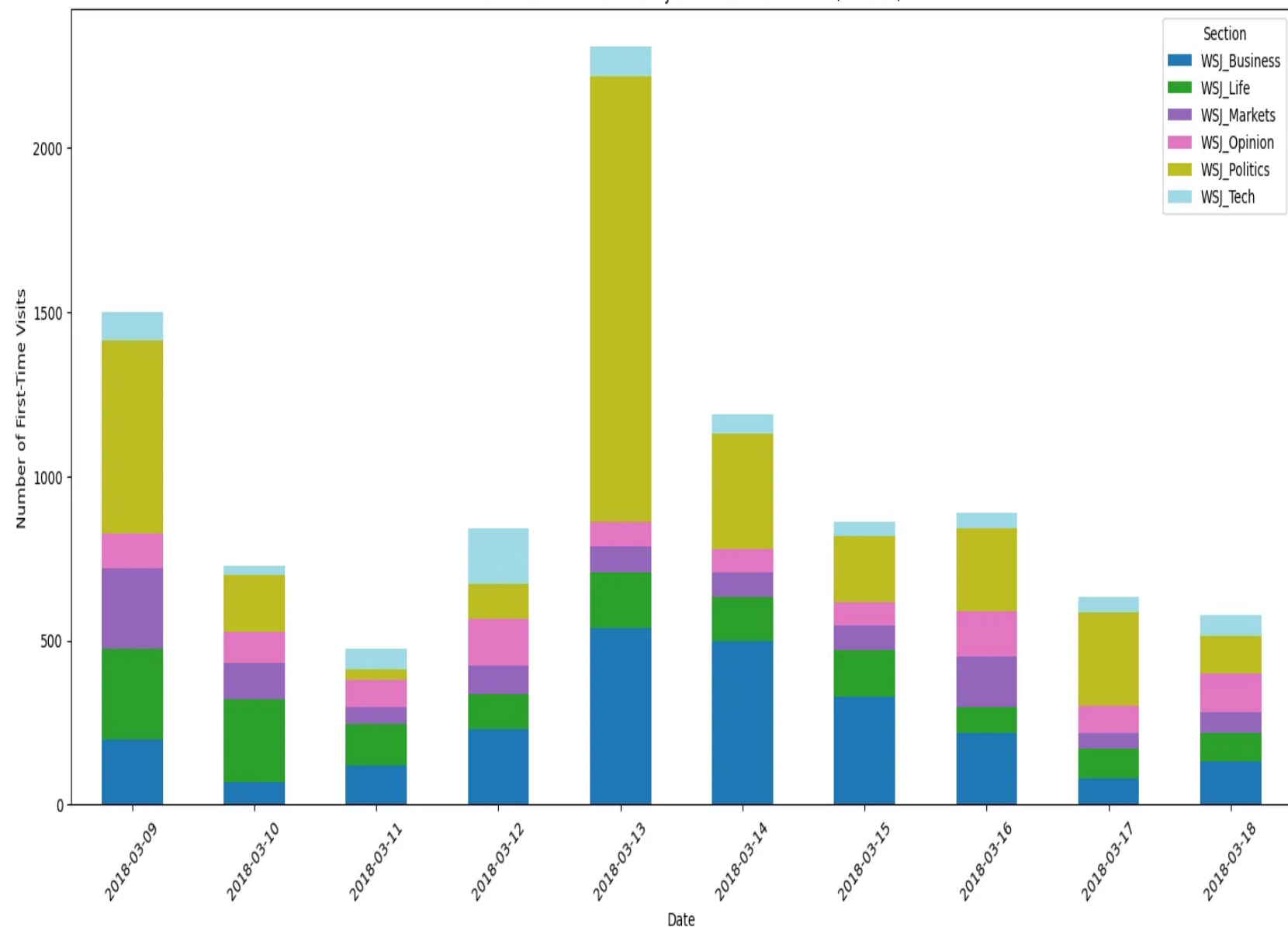
Q5 (4/4)

Question #5: What interesting stories can you tell with this data?

Answer #5.4: The graph shows that the peaks on March 9, 2018, and March 13, 2018, were influenced by concentration on visiting limited number of the most-read articles on those days.

Most Visited Article per Day
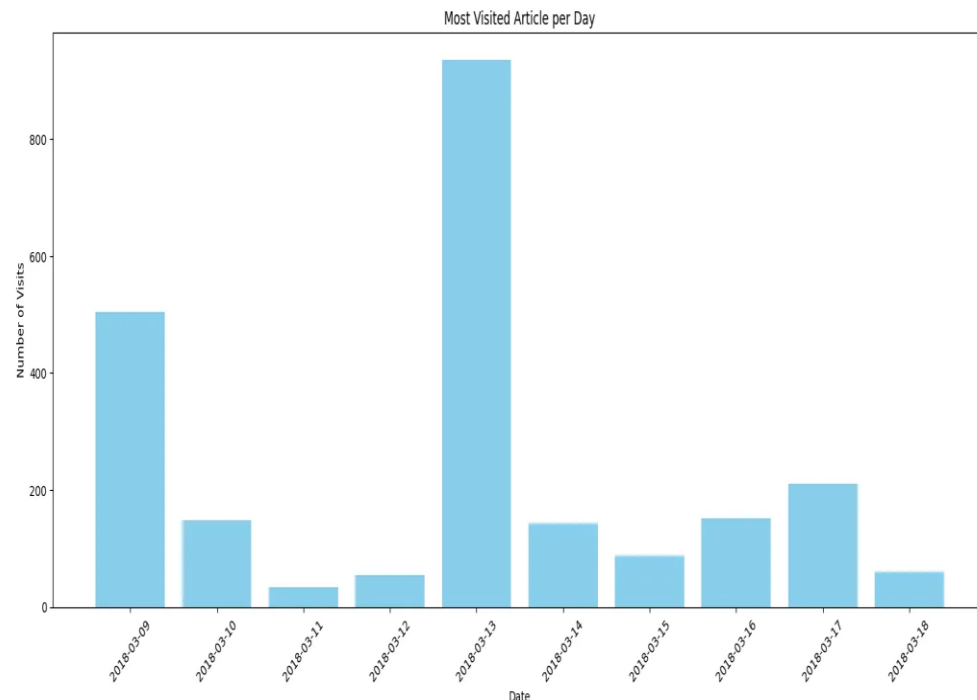
| visitDate	section	headline	visits
2018-03-09	WSJ_Politics	WSJ_Article_Politics and Policy_Trump Lawyers Seek Deal With Mueller to Speed End of Russia Probe	505
2018-03-10	WSJ_Politics	WSJ_Article_Politics and Policy_Trump Lawyers Seek Deal With Mueller to Speed End of Russia Probe	148
2018-03-11	WSJ_Opinion	WSJ_Article_Commentary (U.S.)_The Exhaustion of American Liberalism	34
2018-03-12	WSJ_Opinion	WSJ_Article_Commentary (U.S.)_The Exhaustion of American Liberalism	55
2018-03-13	WSJ_Politics	WSJ_Article_Politics and Policy_Trump's Personal Assistant Is Fired	908
2018-03-14	WSJ_Business	WSJ_Article_Business_Toys 'R' Us Tells Workers It Will Likely Close All U.S. Stores	143
2018-03-15	WSJ_Politics	WSJ_Article_Politics and Policy_Special Counsel Robert Mueller Has Subpoenaed Trump Organization in	88
2018-03-16	WSJ_Politics	WSJ_Article_Politics and Policy_Donald Trump and John Kelly Reach Truce	90
2018-03-17	WSJ_Politics	WSJ_Article_Politics and Policy_Andrew McCabe Kept Notes About Conversations With Trump, Gave Them t	144
2018-03-18	WSJ_Opinion	WSJ_Article_Review & Outlook (U.S.)_The Trump Tariff Layoffs Begin	57

SOLUTION
Part B

QB1

Question #B1: Evolution of number of bridges in categories V. – VII. in Slovakia
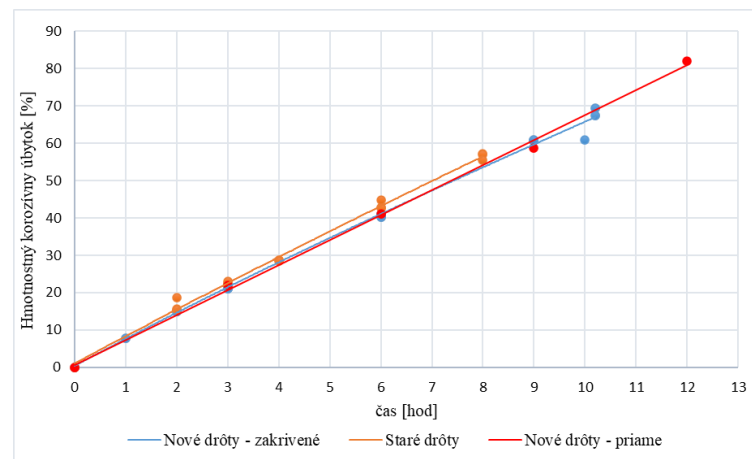
Answer #B1:

QB2
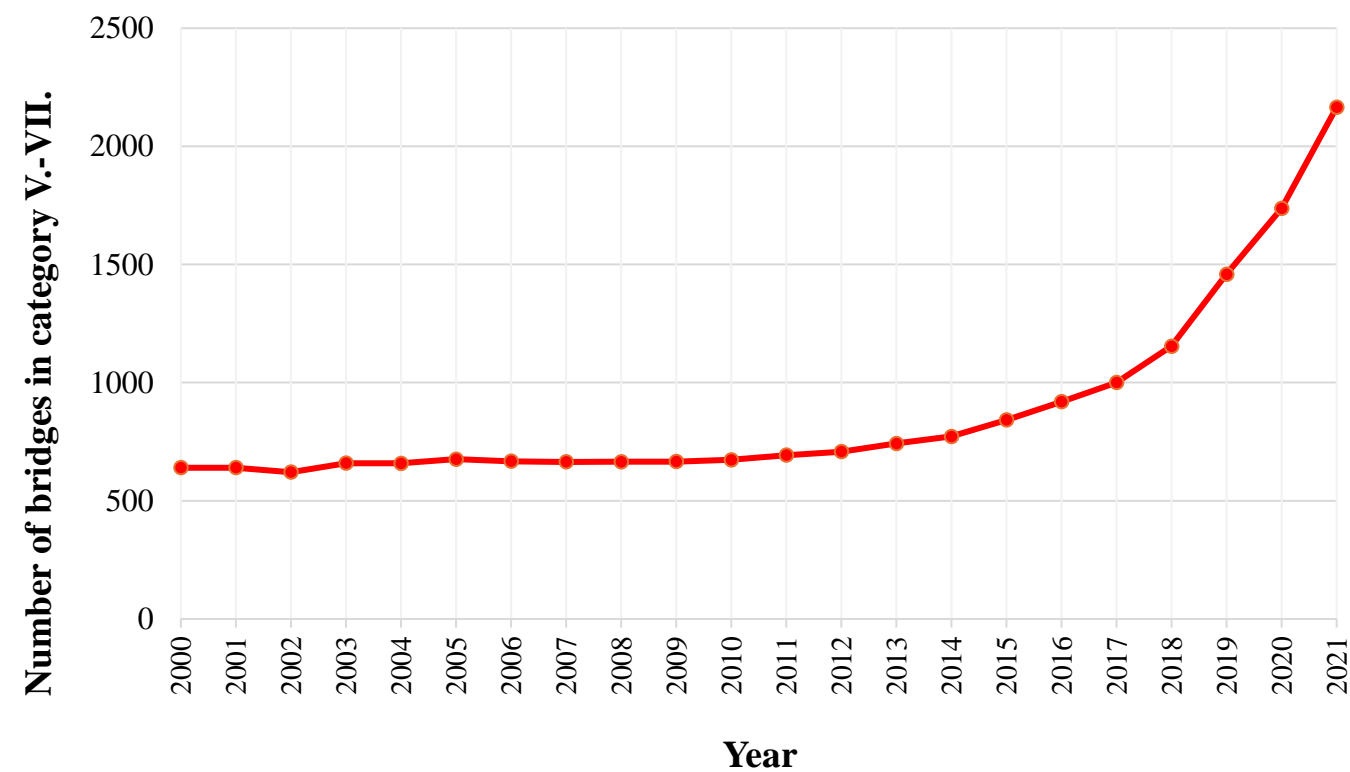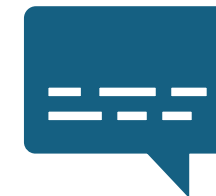
Answer #B2:

Corrosion mass loss versus corrosion time
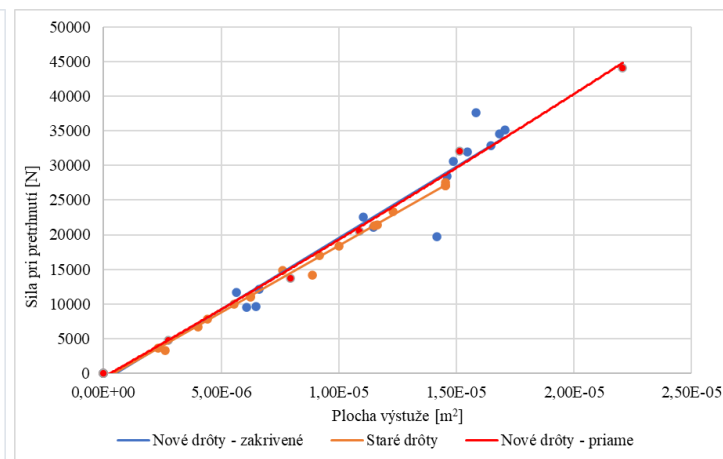
Breaking force to reinforcement area

Question #B2: Results of the primary experiment.
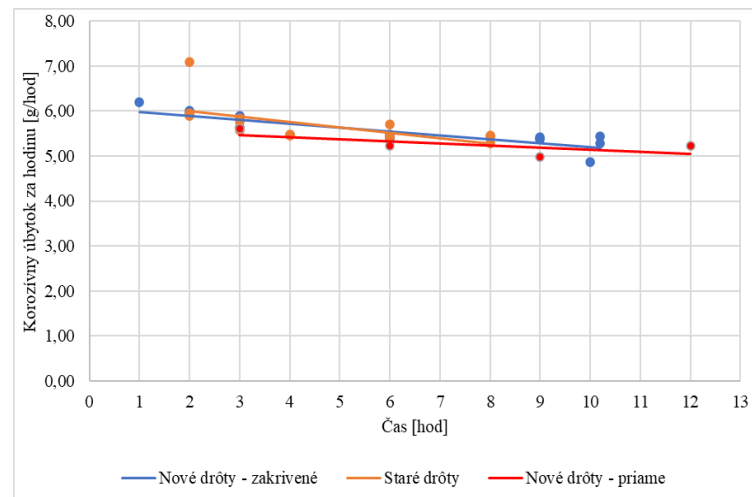
Corrosion rate per hour to the length of the corrosive process

Breaking strength due to corrosive mass loss

QB2

Answer #B2
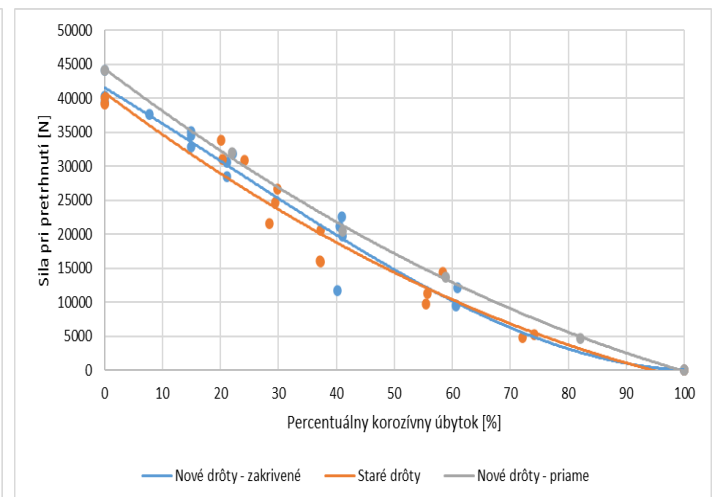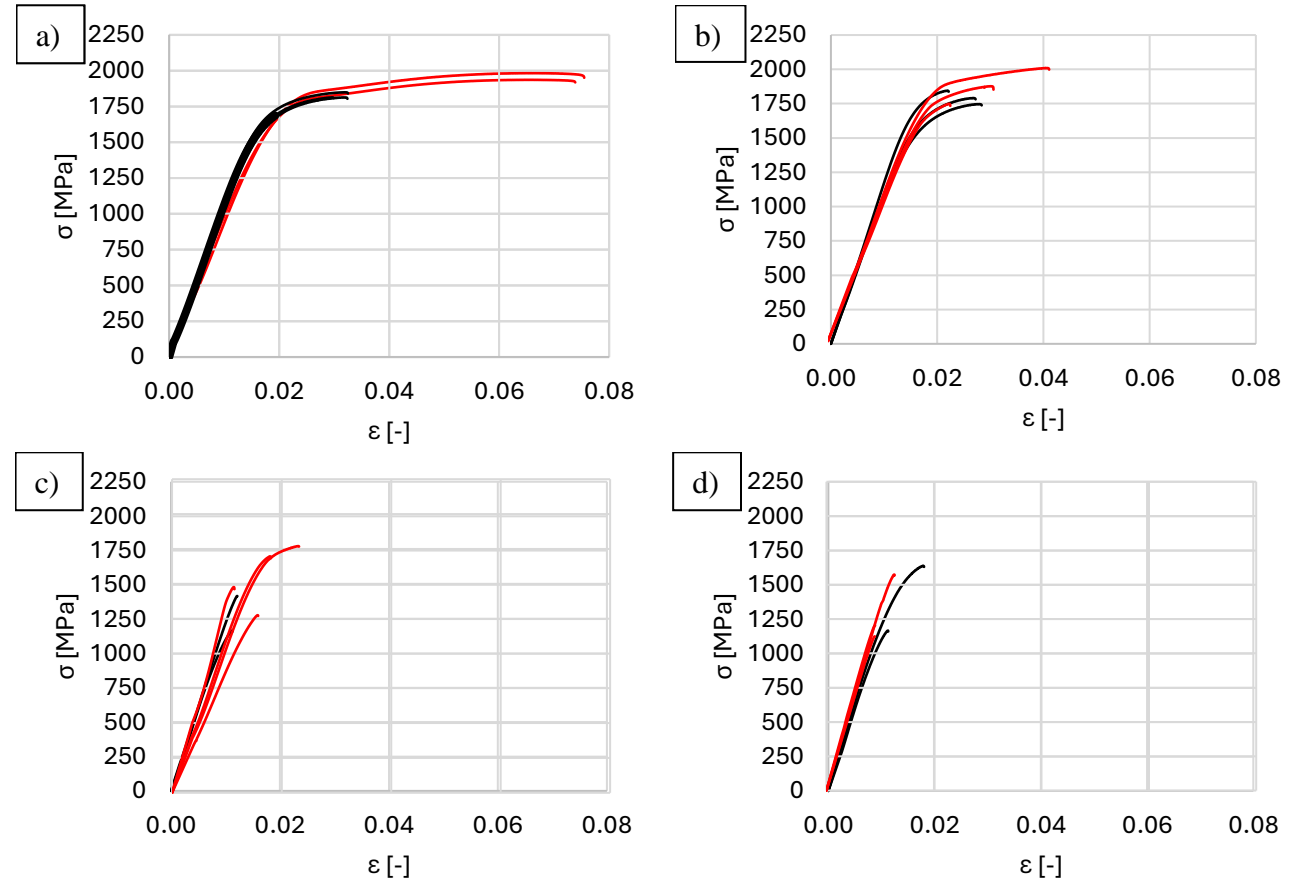

Question #B2: Results of the primary experiment.

Prestressing reinforcement diagram - mass corrosion loss
a)0%, b)20%, c)40%, d)60%

CONCLUSION

FINAL WORDS

This is just a very short part of the projects I have been involved recently.

For more information, feel free to contact me at kopasek.denis@gmail.com