Google Ads Transparency Center Analysis for Advertisers in the European Economic Area (EEA) and Turkey

• • •

Presented By: Melson Heo

Table of Contents

- Data Dictionary
- Introduction
- Project Tools
- Data Cleaning
- Analysis
- Recommendation
- Conclusion

Data Dictionary (Creative Stats)

- Advertiser_id: A unique identifier for a specific advertiser
- Creative_id: A unique identifier for a specific ad
- Creative_page_url: A link to the relevant Ads Transparency Center page for this ad
- Ad_format_type: A description of the ad type
- Advertiser_disclosed_name: The advertiser's name for transparency reporting and disclosures
- Advertiser_legal_name: The legal name of the advertiser
- Advertiser_location: ISO 3166-2 region code (for example "FR" for France)
- Advertiser_verification_status: Whether or not the advertiser has verified their identity with Google yet

Data Dictionary (Creative Stats)

- Region_stats: Collection of regions in which the ad served (in the European Economic Area and Turkey) and their relevant stats
    - Region_code: ISO 3166-2 region code
    - First_shown: YYYY-MM-DD representation of the first date the ad was shown.
    - Last_shown: YYYY-MM-DD representation of the last date the ad was shown\
    - Times_shown_end_date: YYYY-MM-DD representation of the date up to which times_shown
    - Times_shown_lower_bound: The lower bound of the number of times the ad was shown in the associated region.
    - Times_shown_upper_bound: The upper bound of the number of times the ad was shown in the associated region.
    - Times_shown_start_date: YYYY-MM-DD representation of the earliest date for which times_shown is considered (both upper and lower bounds). Impressions are reported for activity since March 1st, 2023.
    - Times_shown_availability_date: YYYY-MM-DD representation of the date on which impressions will become available if no impressions occurred within the reporting period.
    - Surface_serving_stats: Collection of surface which the ad served in and their relevant serving times lower/upper bound.

Data Dictionary (Creative Stats)

- Audience_selection_approach_info: Information on audience selection, broken out by approach type.
  - Demographic_info: Describes if and how demographic info was used to select the ad audience
  - Geo_location: Describes if and how geo location info was used to select the ad audience
  - Contextual_signals: Describes if and how contextual signals were used to select the ad audience
  - Customer_lists: Describes if and how customer lists were used to select the ad audience
  - Topics_of_interest: Describes if and how topics of interest were used to select the ad audience
- Topic: The topic of the ad

Data Dictionary (Removed Creative Stats)

- Creative_page_url: A link to the removed Ads Transparency Center page for this ad
- Region_stats: Collection of regions in which the ad served (in the European Economic Area) and their relevant stats (same as creative_stats)
- Audience_selection_approach_info: Information on audience selection, broken out by approach type (same as creative_stats)
- Disapproval: Information on how, where, and why the ad was removed
    - Removal Reason: The policy/policies that the ad violated
    - Violation Category: The high level policy violation category
    - Use of automated means: Populated if the review of the ad was automated
    - Removal_location: ISO 3166-2 region code
    - Decision_type: Whether the investigation was a result of a Google investigation or a legal notice

Introduction

Introduction

- Advertisements play a pivotal role in promoting products and capturing consumer attention in today's dynamic market landscape.

- However, the world of advertising is as dynamic as it is essential, where campaigns can swiftly transition from gaining immense traction to being abruptly removed for various reasons.

Introducing Google Ads Transparency

- What is Google Ads Transparency?
  - Google Ads Transparency provides users with information about the ads they see, including who paid for them and why they are being shown. This builds trust and empowers users to make informed decisions.
- How can business use this to promote their service?
  - Businesses can leverage Google Ads Transparency to showcase their legitimacy and commitment to ethical advertising. By participating in verification programs and providing accurate information, businesses can build trust with potential customers.

Google Ads Transparency Center

Data Source: Google Ads Transparency

- Data Source: The dataset comprises two tables: creative_stats and removed_creative_stats.
- creative_stats Table:
  - Provides insights on advertisers serving ads in the European Economic Area or Turkey, including their legal name, verification status, disclosed name, and location.
  - Includes ad-specific details such as impression ranges per region (including aggregate impressions for the European Economic Area), first shown and last shown dates, audience selection criteria, ad format, ad topic, and a link to the ad in the Google Ads Transparency Center.
- removed_creative_stats Table:
  - Contains information on ads served in the European Economic Area that were removed by Google, including reasons for removal and per-region information on when they were served.
  - Provides a link to the Google Ads Transparency Center for the removed ad.
- Data Updates: Both tables receive periodic updates, although data may be delayed compared to what appears on the Google Ads Transparency Center website. (Data was extracted from 2/27/2024)

Project Tools

SQL/BigQuery: Extraction/Query

Excel: Analysis

Python/Jupyter Notebook: Data Cleaning/Analysis

Tableau: Visualization/Dashboard

Data Cleaning: creative_stats

- Limited data extraction to 100,000 rows from BigQuery due to extended processing times, ensuring efficient analysis while representing a substantial portion of the dataset.
- Utilized the explode() function to create separate rows for each dictionary entry within the 'region_stats' column, facilitating detailed analysis at the individual dictionary level.
- Employed json_normalize(), concat(), and drop() functions to flatten the dictionaries within the 'region_stats' column, merge the resulting DataFrame with additional data from the 'audience_selection_approach_info' column, and eliminate unnecessary columns, enhancing data clarity and streamlining the dataset for analysis.
- Increased to ~381,000 rows

```python
import pandas as pd

# Explode 'region_stats' column to create separate rows for each dictionary
exploded_results = results.explode('region_stats')

# Normalize the exploded DataFrame to flatten the dictionaries within 'region_stats'
normalized_results = pd.json_normalize(exploded_results['region_stats'])

# Create a new DataFrame by converting the 'audience_selection_approach_info' column to a DataFrame
audience_info_df = pd.DataFrame(results['audience_selection_approach_info'].tolist())

# Concatenate the original DataFrame with the normalized DataFrame and audience_info_df
merged_results = pd.concat([exploded_results.reset_index(drop=True), normalized_results.reset_index(drop=True), audience_info_df.reset_index(drop=True)], axis=1)

# Drop the original 'region_stats' and 'audience_selection_approach_info' columns
merged_results = merged_results.drop(columns=['region_stats', 'audience_selection_approach_info'])

# Display the merged results
print("Merged Results:")
merged_results.head()
```

Data Cleaning: removed_creative_stats

- Limited data extraction to 100,000 rows from BigQuery due to extended processing times, ensuring efficient analysis while representing a substantial portion of the dataset.
- Utilized explode() function to separate rows for each dictionary entry in 'region_stats', enhancing data structure.
- Employed explode() function for 'disapproval' column, increasing data granularity.
- Converted 'audience_selection_approach_info' column to DataFrame, creating separate columns.
- Dropped unnecessary columns ('region_stats', 'disapproval', 'audience_selection_approach_info') for streamlined analysis.
- Increase to ~680,000 rows as a result

```python
import pandas as pd

# Explode 'region_stats' column to create separate rows for each dictionary
exploded_results = results.explode('region_stats')

# Normalize the exploded DataFrame to flatten the dictionaries within 'region_stats'
normalized_region_stats = pd.json_normalize(exploded_results['region_stats'])

# Concatenate the original DataFrame with the normalized DataFrame for 'region_stats'
merged_results_region = pd.concat([exploded_results.reset_index(drop=True), normalized_region_stats.reset_index(drop=True)], axis=1)

# Drop the 'region_stats' column
merged_results_region = merged_results_region.drop(columns=['region_stats'])

# Explode 'disapproval' column to create separate rows for each dictionary
exploded_disapproval = merged_results_region.explode('disapproval')

# Normalize the exploded DataFrame to flatten the dictionaries within 'disapproval'
normalized_disapproval = pd.json_normalize(exploded_disapproval['disapproval'])

# Concatenate the DataFrame with the normalized DataFrame for 'disapproval'
merged_results_disapproval = pd.concat([exploded_disapproval.reset_index(drop=True), normalized_disapproval.reset_index(drop=True)], axis=1)

# Drop the 'disapproval' column
merged_results_disapproval = merged_results_disapproval.drop(columns=['disapproval'])

# Convert 'audience_selection_approach_info' column to a DataFrame
audience_info_df = pd.DataFrame(merged_results_disapproval['audience_selection_approach_info'].tolist())

# Concatenate the original DataFrame with the new DataFrame containing separated columns for 'audience_selection_approach_info'
merged_results_audience = pd.concat([merged_results_disapproval.reset_index(drop=True), audience_info_df.reset_index(drop=True)], axis=1)

# Drop the 'audience_selection_approach_info' column
merged_results_final = merged_results_audience.drop(columns=['audience_selection_approach_info'])

# Displaying the final DataFrame
merged_results_final.head()
```

Analysis

Understanding Impressions

- Impressions measure how many times your ad is displayed, providing a basic understanding of your potential reach. This can be valuable for brand awareness campaigns aiming to get their message seen by a broad audience.
- Impressions allow you to compare campaigns and track their performance over time. This can help identify trends and assess the relative visibility of different ad strategies.
- Defining impression goals can be relevant for specific campaign objectives, such as reaching a minimum number of potential customers or achieving a certain level of brand exposure.
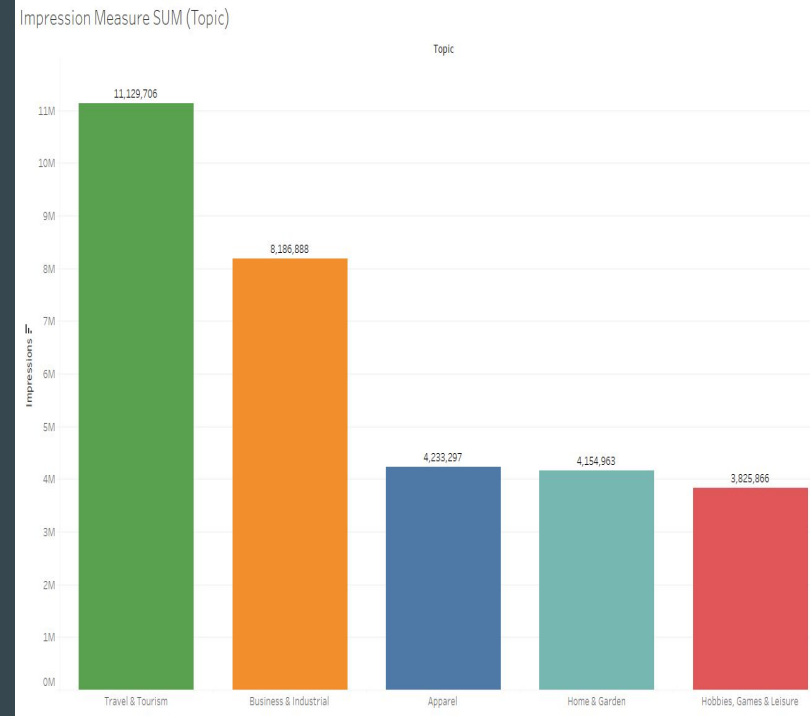
Impressions

- Impressions measure the total count of times an ad is fetched and displayed on users' screens, indicating the potential reach and visibility of digital advertising campaigns.
- Impression: Times Shown End Date - Times Shown Start Date (within creative_stats)
- It will be measured based on days
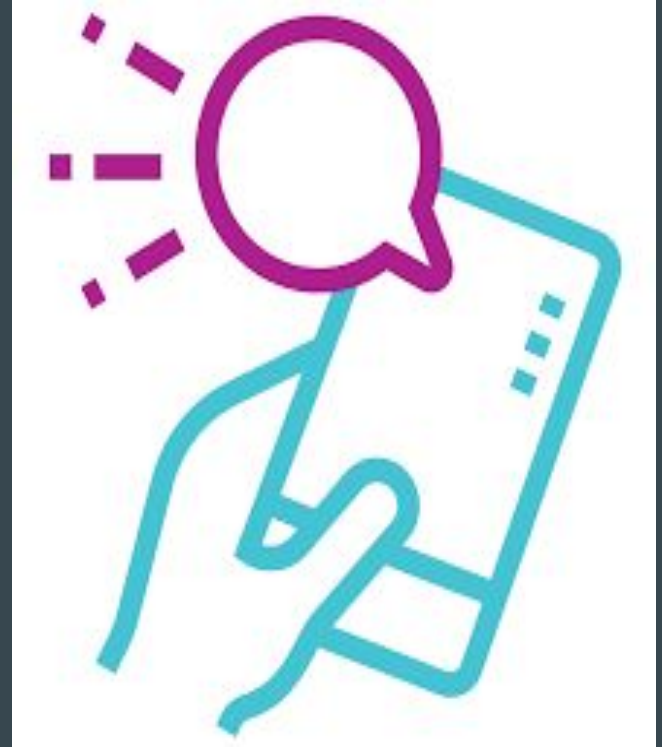
Impression Measure (Topic)

- In this dataset, there were 51 topics. To focus a specific demographic, it was better to focus the top 5 best performing topics in terms of SUM.
- In this first graph, this shows the top 5 best performing topics in SUM. This shows that many of the ads included Travel, Business, and Hobbies.

Impression Measure SUM (Topic)

Topic

11,129,706

8,186,888

4,233,297   4,154,963   3,825,866

Impressions

Travel & Tourism   Business & Industrial   Apparel   Home & Garden   Hobbies, Games & Leisure

Impression Measure (Topic)

- Although we can see the SUM of Topics can be skewed due to the amount of ads within the year, we should also check the AVG Impressions.
- As within the AVG Impressions, we can see how each impressions has skewed more for niche sides of things.

Impression Measure AVG (Topic)
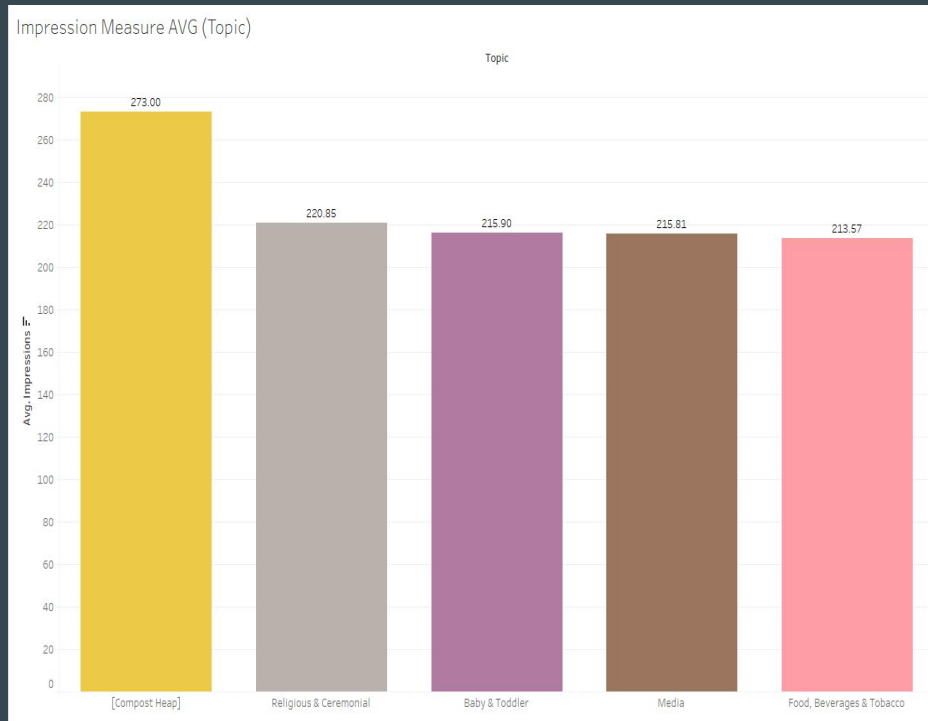
Impression Measure (Ad Format)

- Ad formats: Text, Video, and Image
- Text Ads:
  - Consist primarily of text and commonly seen on search engines or websites.
  - Include a headline, description, and link to the advertiser's website.
- Image Ads:
  - Incorporate visual elements like pictures or graphics along with text.
  - Displayed on websites, social media platforms, or in digital formats.
- Video Ads:
  - Utilize video content to convey a message, ranging from short clips to longer-form content.
  - Displayed on online video platforms, social media, or streaming services, and highly engaging for storytelling and brand building.

Impression Measure (Ad Format)

- Text ads have the highest average impressions, highlighting their effectiveness in engaging audiences.
- Video ads show the lowest average impressions, indicating potential for optimization to increase reach.
- The prevalence of text ads suggests a strong preference among advertisers for this format due to its consistent performance in generating impressions.

Impression Measure (Ad Format)
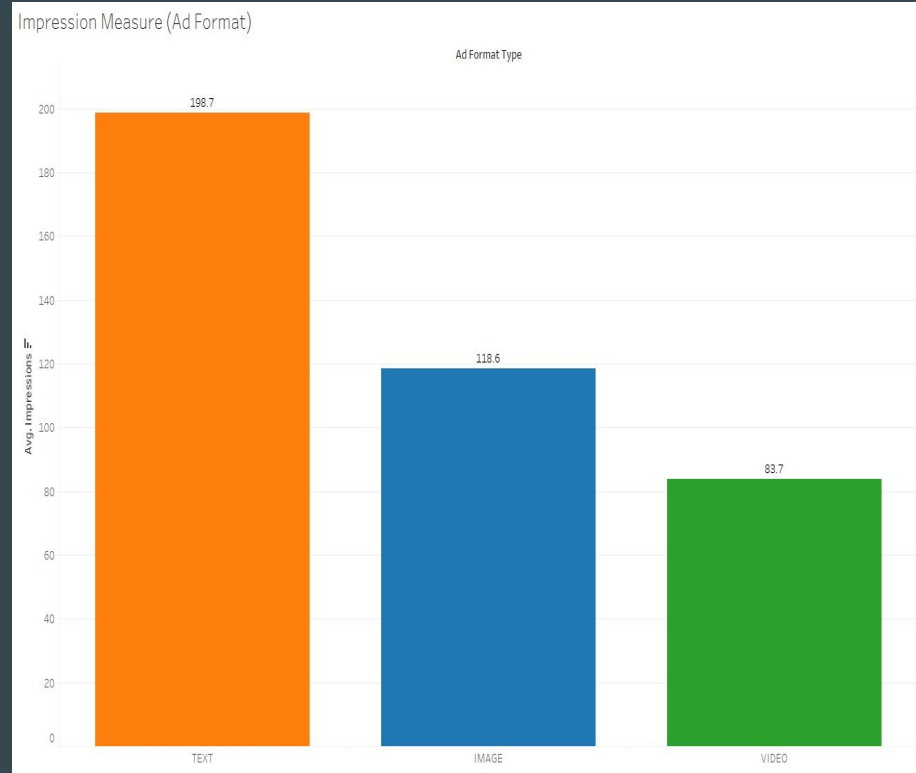
Impression Ad Format First Shown

Impression Ad Format Last Shown
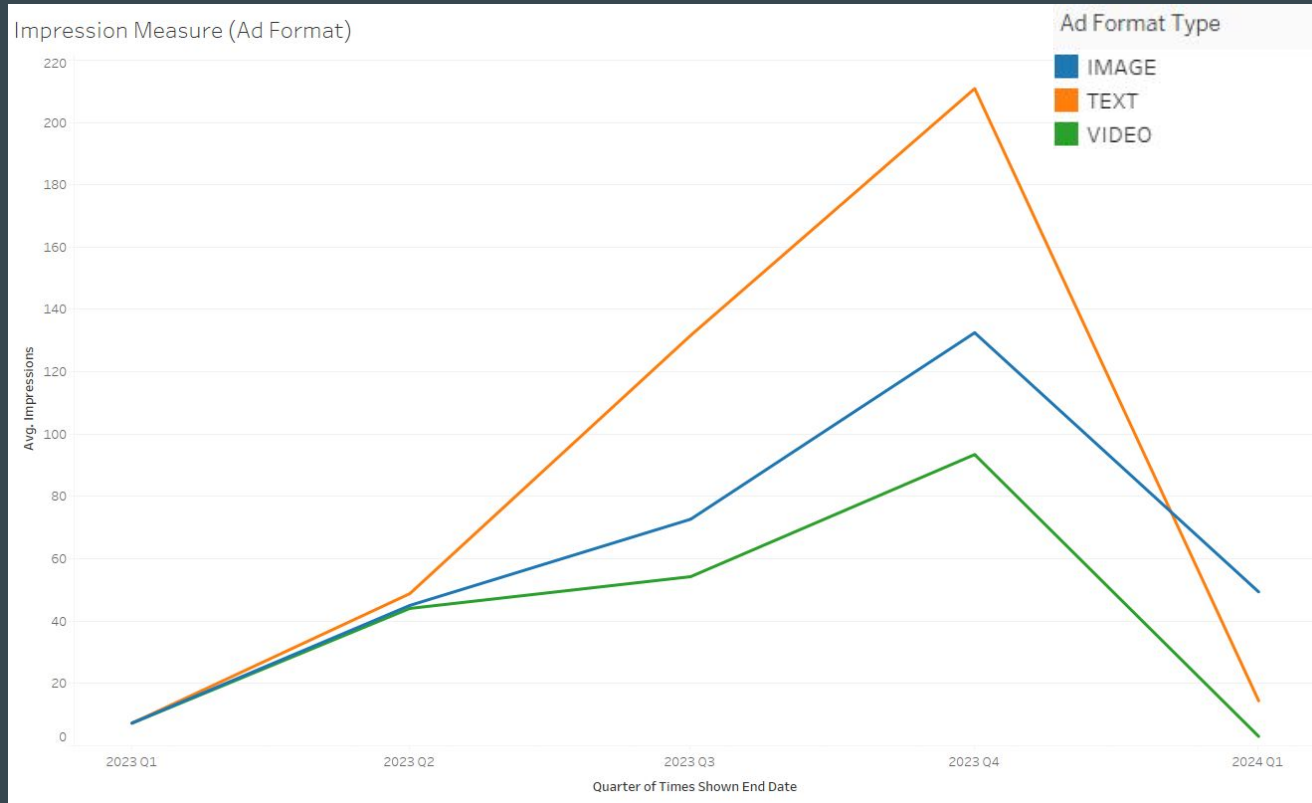
Tackling Removed Ads

- While impressions tell you how many people saw your ad, ad removal data reveals why they didn't see it. This deeper understanding helps identify areas for improvement.
- Analyzing ad removal reasons goes beyond basic metrics. It uncovers issues like irrelevant targeting, misleading content, or technical glitches impacting campaign effectiveness.
- Understanding removal reasons allows you to optimize your campaigns. You can refine ad content, adjust targeting, or address technical issues to maximize your reach and impact.

Violation Category Reasons

- There are four categories: Prohibited Practices, Restricted Content and Features, Editorial and Technical, and Prohibited Content
- Within the violation category, we can see that prohibited practices have the highest percentage.
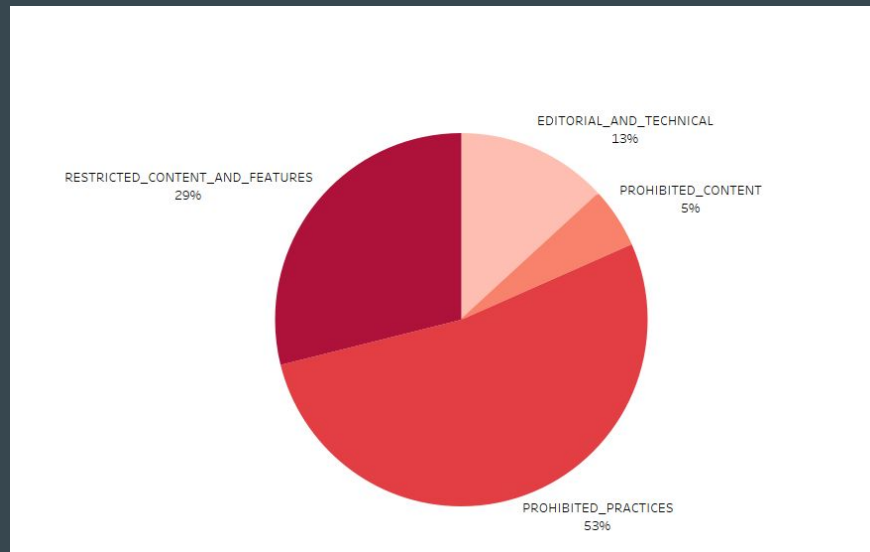
Ad Removal Breakdown

- Throughout the violation category, we can see that the violation have their own separate ad removal reasons
- We can break down each of the violation category into ad removal reasons
- This reveals that account suspension is a major reason for ad removal. This suggests a prevalence of policy violations, highlighting the importance of adhering to advertising guidelines.

Violation Category
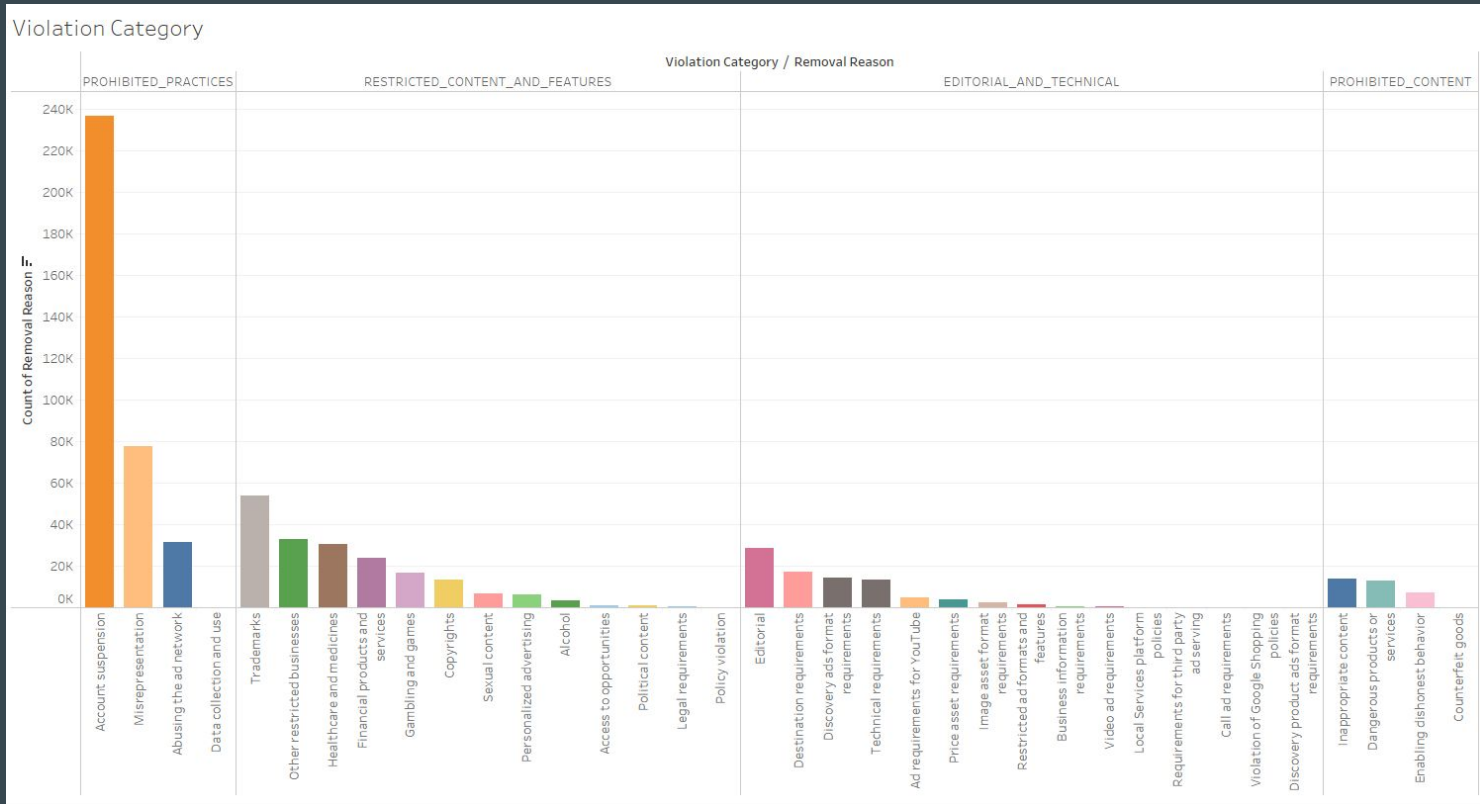
Region Comparison

- By analyzing both impressions (ad views) and ad removals, we can pinpoint the top 10 regions generating the most ad views. This helps us understand where our potential customer base is located and tailor strategies accordingly.
- Comparing impressions with ad removals reveals regions with high ad viewership but also frequent removals. This allows us to investigate potential issues like irrelevant targeting or policy violations, ensuring your message reaches the right audience effectively.
- By analyzing these insights, we can refine Google Ads campaigns for better targeting and improved performance.
- While the initial data included a significant number of entries from the EEA (European Economic Area), it lacked specific regional details. To enable a more granular comparison across regions, we applied filters to focus on specific areas outside the broader EEA category.

Top 10 Regions with Highest Average Impressions
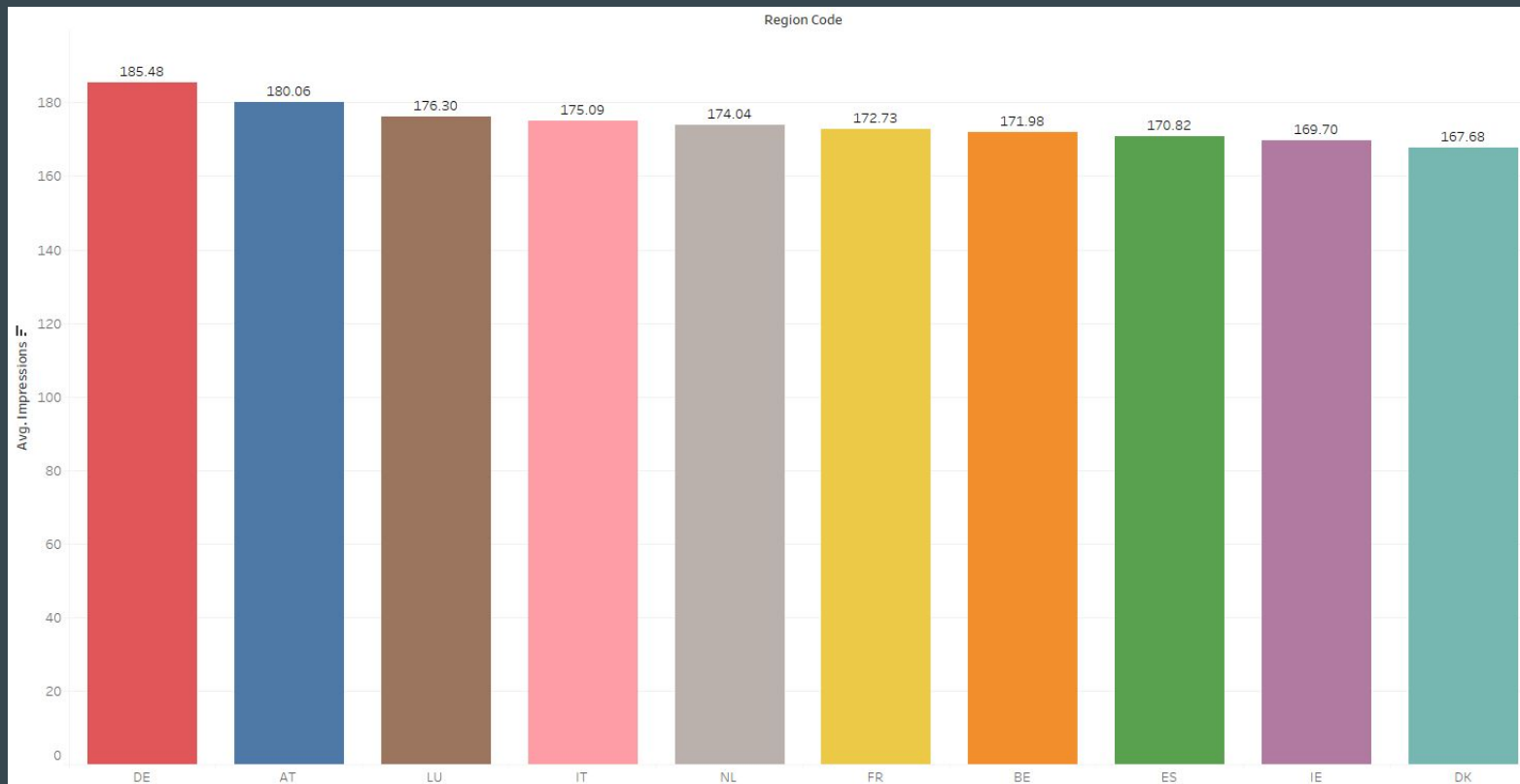
Top 10 Regions with Highest Violation Category Count

Region Code

Count of Violation Category

41,820	35,538	32,119	31,831	28,684	27,310	22,645	21,046	20,120	19,651
DE	FR	ES	IT	PL	NL	RO	BE	HU	PT
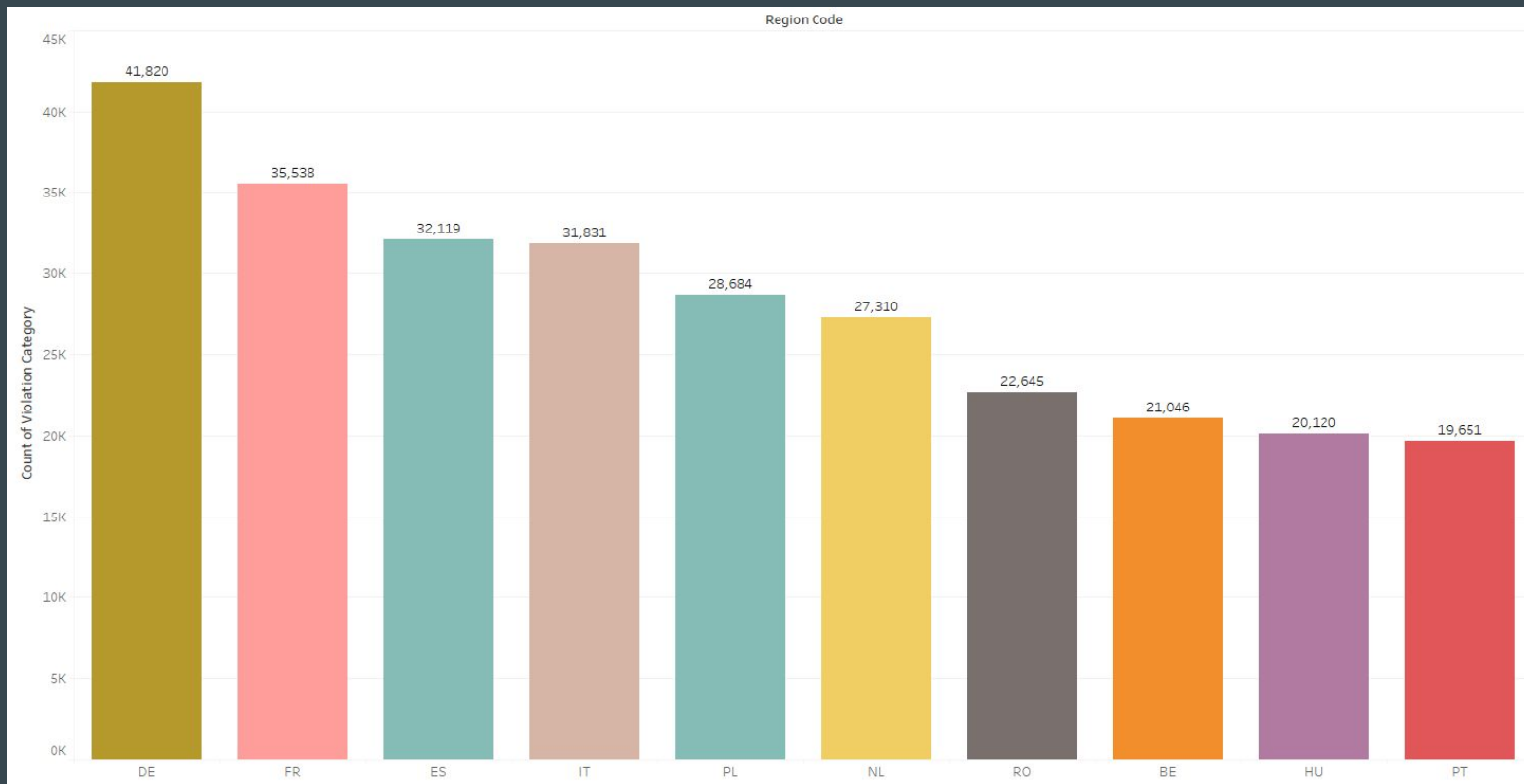
Audience Criteria

- We can track customers behavior by using these five metrics:
  - Demographic_info: Describes if and how demographic info was used to select the ad audience
  - Geo_location: Describes if and how geo location info was used to select the ad audience
  - Contextual_signals: Describes if and how contextual signals were used to select the ad audience
  - Customer_lists: Describes if and how customer lists were used to select the ad audience
  - Topics_of_interest: Describes if and how topics of interest were used to select the ad audience
- Four Entries:
  - CRITERIA_INCLUDED
  - CRITERIA_INCLUDED_AND_EXCLDUED
  - CRITERIA_UNUSED
  - CRITERIA_EXCLUDED

Audience Demographic (Creative)

Demographic Info / Topics Of Interest

Contextual Signals	Geo Location	Customer Lists	Null	CRITERIA_EXCLUDED				CRITERIA_INCLUDED				CRITERIA_INCLUDED_AND_EXCLUDED				CRITERIA_UNUSED			
			Null	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..	CRITERI..
Null	Null	Null	0																
CRITERIA_EXCLUDED	CRITERIA_INCLUDED	CRITERIA_INCLUDED									1								
CRITERIA_INCLUDED	CRITERIA_EXCLUDED	CRITERIA_EXCLUDED							4		137		22		16				
		CRITERIA_INCLUDED							21		43		6		4				2
		CRITERIA_INCLUDED_A..							1		10		7						
		CRITERIA_UNUSED							41		442		5		8				6
	CRITERIA_INCLUDED	CRITERIA_EXCLUDED						10	44	1	433		41	4	48				1
		CRITERIA_INCLUDED			1		1	1	1,630	1	1,866		273	1	163		15	1	133
		CRITERIA_INCLUDED_A..			2	1			95		303	1	60		53		1		4
		CRITERIA_UNUSED			3		11	6	854	2	14,185	3	390		559		11		272
	CRITERIA_INCLUDED_A..	CRITERIA_EXCLUDED						1	13	1	364	1	15	1	26				2
		CRITERIA_INCLUDED						1	321	6	393		40		43		3		16
		CRITERIA_INCLUDED_A..							36	5	113		21	1	15		4		3
		CRITERIA_UNUSED						4	127		3,360	2	28		108				66
	CRITERIA_UNUSED	CRITERIA_EXCLUDED									28								
		CRITERIA_INCLUDED							48		88		7		4		2		17
		CRITERIA_INCLUDED_A..							2		7				3				1
		CRITERIA_UNUSED						1	124		830		5		15				29
CRITERIA_INCLUDE..	CRITERIA_EXCLUDED	CRITERIA_EXCLUDED						1	3		72	1	8		22				3
		CRITERIA_INCLUDED							43		119		3		4			1	10
		CRITERIA_INCLUDED_A..					1		22		152	1	7		6			1	6
		CRITERIA_UNUSED			1		2	1	29		866	2	4		17				7
	CRITERIA_INCLUDED	CRITERIA_EXCLUDED			2		10	37	200	2	1,973	10	159	1	294		2		25
		CRITERIA_INCLUDED			42	1	32	8	2,900	15	6,849	2	455	13	464		123	1	372
		CRITERIA_INCLUDED_A..		2	59	2	65	5	480	4	1,052	5	194	6	247	1	94	1	259
		CRITERIA_UNUSED		1	82	3	79	78	1,070	13	28,853	49	783	13	1,659		80	2	932
	CRITERIA_INCLUDED_A..	CRITERIA_EXCLUDED					17	26	126	7	1,187	11	73	2	239	1	1		4
		CRITERIA_INCLUDED			1		2	4	1,029	5	1,843	4	137	5	176		20		49
		CRITERIA_INCLUDED_A..			15		9	6	338	13	553	8	103	5	128	1	30	4	44
		CRITERIA_UNUSED			9		12	33	516	7	9,926	38	188	8	655		9	1	119
	CRITERIA_UNUSED	CRITERIA_EXCLUDED							2		14				1				
		CRITERIA_INCLUDED			1				33	1	148		5	1	6		1		17
		CRITERIA_INCLUDED_A..							4		206		2		1		1	1	18
		CRITERIA_UNUSED			1		5		24		852		8		13		2		110
CRITERIA_UNUSED	CRITERIA_INCLUDED	CRITERIA_INCLUDED									19				2				1
		CRITERIA_UNUSED									86				3				
	CRITERIA_UNUSED	CRITERIA_UNUSED							1		46								408

Audience Demographic (Removed)

Customer Lists / Geo Location — column groups: CRITERIA_EXCLUDED, CRITERIA_INCLUDED, CRITERIA_INCLUDED_AND_EXCLUDED, CRITERIA_UNUSED (each with 4 "CRITERI.." sub-columns).

Contextual Signals	Demographic Info	Topics Of Interest	EXCL 1	EXCL 2	EXCL 3	EXCL 4	INCL 1	INCL 2	INCL 3	INCL 4	INC&EXC 1	INC&EXC 2	INC&EXC 3	INC&EXC 4	UNUSED 1	UNUSED 2	UNUSED 3	UNUSED 4
CRITERIA_EXCLUDED	CRITERIA_INCLUDED	CRITERIA_UNUSED													20		12	37
CRITERIA_INCLUDED	CRITERIA_EXCLUDED	CRITERIA_INCLUDED						2								12		
		CRITERIA_UNUSED						4				88				7	2	
	CRITERIA_INCLUDED	CRITERIA_EXCLUDED		7												10	27	11
		CRITERIA_INCLUDED		220	26	17	545	9,885	1,674	1,835	139	572	163	56	901	13,056	942	1,862
		CRITERIA_INCLUDED_A..																3
		CRITERIA_UNUSED	210	3,827	527	33	1,972	7,595	1,473	1,647	103	614	239	189	11,558	93,763	14,411	48,686
	CRITERIA_INCLUDED_A..	CRITERIA_EXCLUDED						20							32	14	78	
		CRITERIA_INCLUDED	101	119	52		276	4,968	566	239	148	316	78	62	524	8,936	435	747
		CRITERIA_UNUSED	152	198	144		211	1,983	465	144		312	132	27	657	11,671	1,481	767
	CRITERIA_UNUSED	CRITERIA_INCLUDED						49	2			16	2			17	2	
		CRITERIA_UNUSED		6				14		8	2					651	2	871
CRITERIA_INCLUDED_A..	CRITERIA_EXCLUDED	CRITERIA_INCLUDED						49				36	44			93	13	
		CRITERIA_UNUSED		10	9			31			13	69	14	3		184	10	
	CRITERIA_INCLUDED	CRITERIA_EXCLUDED		64	140			2	2				2		14	189	376	
		CRITERIA_INCLUDED	32	370	158		469	7,323	3,041	406	1,973	1,340	2,877	2	3,937	7,705	2,414	1,206
		CRITERIA_INCLUDED_A..		2	8			32	14				4		61	21	258	2
		CRITERIA_UNUSED	1,243	5,105	4,144	202	2,712	18,477	3,265	1,130	667	3,399	2,066	1,086	46,592	120,864	40,358	24,883
	CRITERIA_INCLUDED_A..	CRITERIA_EXCLUDED		21	22			2							98	1,059	894	
		CRITERIA_INCLUDED	122	368	175	96	80	2,117	368	108	31	330	142	29	759	8,495	1,265	169
		CRITERIA_INCLUDED_A..			2				28			18	60		23	166	142	
		CRITERIA_UNUSED	441	987	1,081	20	391	1,807	367	128	320	824	609		5,231	38,455	6,023	1,489
	CRITERIA_UNUSED	CRITERIA_INCLUDED			35			318	60	2		157	4			33	19	14
		CRITERIA_UNUSED		106	135		58	1,194	138	51	2	195	12	3	2,347	3,496	565	1,136
CRITERIA_UNUSED	CRITERIA_INCLUDED	CRITERIA_UNUSED		4				39								1,413		1,058
	CRITERIA_INCLUDED_A..	CRITERIA_UNUSED						64								70		8
	CRITERIA_UNUSED	CRITERIA_UNUSED																58

Comparing Audience Demographic

- Both demonstrate an identical peak in audience demographic distribution, highlighting a consistent trend across the analyzed data sets.
- Leveraging demographic insights and geographical data, advertisements employ targeted strategies to effectively engage with and track audience behavior.
- The alignment in the highest demographic representation shows the strategic utilization of demographic and geo-location data to optimize audience targeting and engagement in advertising campaigns.

Contextual Signals: CRITERIA_INCLUDED_AND_EXCLUDED
Customer Lists: CRITERIA_UNUSED
Demographic Info: CRITERIA_INCLUDED
Geo Location: CRITERIA_INCLUDED
Topics Of Interest: CRITERIA_UNUSED
Count of Topics Of Interest: 28,853

Contextual Signals: CRITERIA_INCLUDED_AND_EXCLUDED
Customer Lists: CRITERIA_UNUSED
Demographic Info: CRITERIA_INCLUDED
Geo Location: CRITERIA_INCLUDED
Topics Of Interest: CRITERIA_UNUSED
Count of Creative Page Url: 120,864

Recommendations and Conclusion

Recommendations

- Optimize ad campaigns by prioritizing personalized messaging tailored to individual needs rather than targeting general demographics, aiming to achieve the highest average impression rates.
- Consider focusing advertising efforts on text ads format, which consistently yields the highest average impressions, maximizing the potential reach and visibility of ad campaigns.
- Mitigate the risk of account suspension, which can adversely impact credibility and campaign effectiveness, by implementing proactive measures to ensure compliance with advertising policies and guidelines.
- For initial advertising endeavors, prioritize targeting specific demographics such as DE (Germany) and FR (France), strategically allocating resources to optimize ad performance and maximize ROI.
- Enhance ad targeting precision and effectiveness by leveraging demographic information and geo-location data, enabling more targeted and impactful advertising campaigns tailored to the unique characteristics and preferences of the audience.

Conclusion

- In conclusion, the findings from the Google Ads Transparency Center underscore the imperative for ongoing vigilance and strategic adaptation in response to the dynamic shifts in consumer behavior and advertising trends.

Sources

[1] BigQuery Public Database

[2] https://support.google.com/adspolicy/answer/2375414?hl=en

[3] https://www.bigcommerce.com/glossary/digital-marketing-impressions/

[4] https://www.singlegrain.com/blog/ms/google-ads-transparency/

[5] https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Country_codes

Thank You!