London School of Economics

Data Analytics Career Accelerator

Data Analytics for Business

2Market Business Insights

Exploratory Analysis and Presenting Insights

Word Count: 948

Date: 18/Dec/2023

Business Analysis Report: Unlocking Growth for 2Market in a Post-Financial Crisis Landscape

Introduction: In the challenging aftermath of the 2008 financial crisis, 2Market faced a substantial sales downturn during Q3 2013. This report delves into customer demographics, spending behaviour, and marketing channels to rejuvenate global sales. The central questions guiding this analysis are: How can we re-engage inactive customer segments. What are specific factors creating or contributing to non-spending? To what degree are our existing marketing channels failing to effectively engage these specific customer groups? How can we ensure customer retention, as well as engagement of new customers?

Target audience for this presentation is the Marketing Strategy Team at 2Market. This includes Marketing Managers, Data Analysts, and Advertising Specialists responsible for shaping and executing the company's advertising strategies. The presentation aims to equip them with actionable insights into customer spending patterns, advertising channel effectiveness, and key demographic trends. By providing detailed analytics and interactive visualisations, the Marketing Strategy Team can make informed decisions, optimise campaigns, and tailor strategies to effectively reach target audiences, thereby enhancing the overall impact of 2Market's marketing initiatives.

Analytical Approach: Data cleaning process in Excel, ensuring data integrity. Outliers, identified through descriptive statistics, were removed for a cleaner dataset. PivotTables utilised for demographic trends, illustrating customer characteristics.

- Spellcheck and check for ‘blanks’: no errors.
- Conditional formatting applied to check for duplicates: none found.
- ‘2n cycle’ renamed ‘Master’ for coherence.
- ‘Dt_customer’ dates corrected:

Figure xx

=TEXTSPLIT(TEXT(B2,"dd/mm/yyyy"),"/")

=DATE function, to combine and standardise date format

=IF(L2<DATE(2000,1,1), L2 + DATE(100, 0, 0), L2)

- Income column: $ symbol removed with Find & Replace. Converted to ‘currency’.
- Data Types: each column checked and adjusted where necessary.
- Customer Age calculated with Birth Year and ‘Dt_Customer’ (assuming dataset is from 2014).
- Clear definitions could not be determined, transferred to different category for coherence.

Figure xx

‘YOLO’, **‘Absurd’**, and **‘Alone’** renamed to **‘Single’**

=IF(OR(E2="Alone",E2="Absurd",E2="YOLO"),"Single",E2)

Figure showing Scatterplot and Outliers

Figure showing comparison of descriptive statistics before and after removal of $666k outlier

- Inquiry with Descriptive Statistics revealed significant right-skew, indicating few individuals in high income bracket.

Income			*Income (£666k removed)*	
Mean	52247.25		Mean	51969.86
Standard Error	534.7508		Standard Error	457.3861
Median	51381.5		Median	51373
Mode	7500		Mode	7500
Standard Deviation	25173.08		Standard Deviation	21526.32
Sample Variance	6.34E+08		Sample Variance	4.63E+08
Kurtosis	159.6367		Kurtosis	0.713549
Skewness	6.763487		Skewness	0.34735
Range	664936		Range	160667
Minimum	1730		Minimum	1730
Maximum	666666		Maximum	162397
Sum	1.16E+08		Sum	1.15E+08
Count	2216		Count	2215

- Further adjustments made through IQR query. Outliers removed accordingly.

AGE

QTL 1	36
QTL 3	54
IQR	18
Lower Limit	9
Upper Limit	81

INCOME

QTL 1	35303
QTL 3	68522
IQR	33219
Lower Limit	-14525.5
Upper Limit	118350.5

- Outliers removed = Age ID: 11004, 1150, 7829; Income ID: 9432, 1503, 1501, 5336, 8475, 4931, 11181, 5555 (highlighted in original (raw) dataset for future reference – if needed).
- PivotTables were utilised to investigate demographic trends.

Excel Analysis Highlights: Detailed Excel analyses involved correcting date formats, converting currency symbols, and adjusting data types. Outliers were identified through a Scatterplot and further refined using an IQR query. Renaming columns for clarity and aligning data with SQL headers streamlined the dataset for further analysis.

Key Findings and Outcomes: Demographic insights revealed that the average customer age is 44, with an income of $51,622 and average spending of $606.82. Most common demographics include customers with graduation and master's degrees, married individuals, and those from Spain and South Africa. Notably, Montenegro posed challenges due to limited data.

Tableau and PostgreSQL Analysis

Spend-Discount correlation: noticeable trend is that total spend increases with an increase in the number of discount deals up until around 5 or 6, after which it begins to decline. The data points are more densely packed at lower numbers of discount deals. Forward to sales and marketing to ensure appropriate discounts are selected.

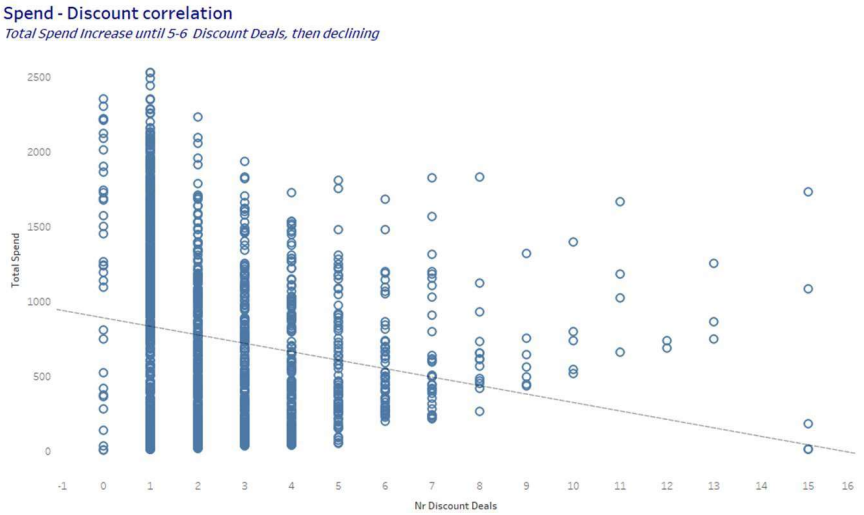


Figure showing best selling products per country 'alcohol and meat'

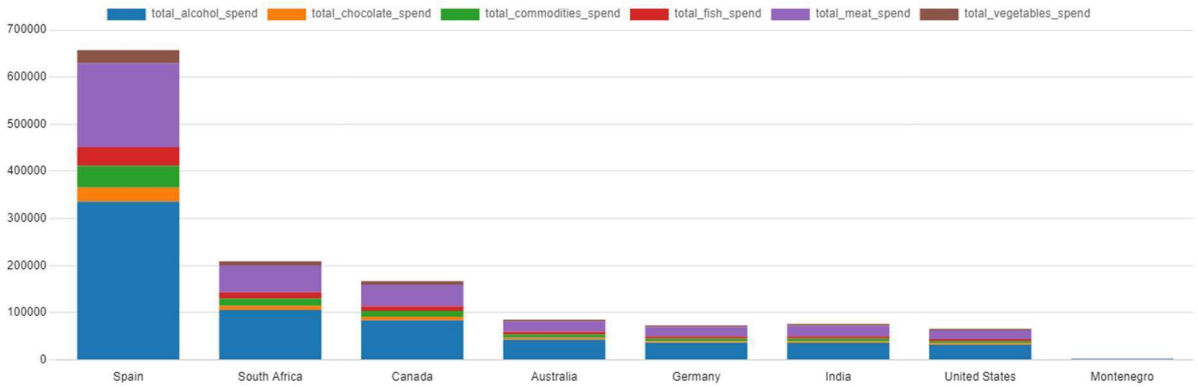
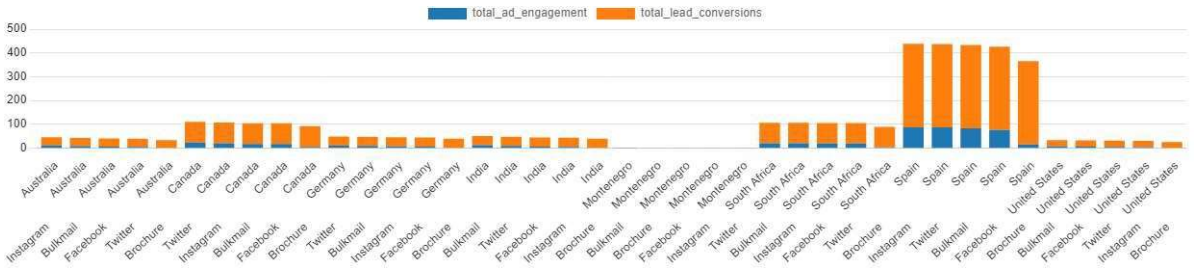


Figure showing Total Ad Engagement and Lead Conversions by Country and Ad channel



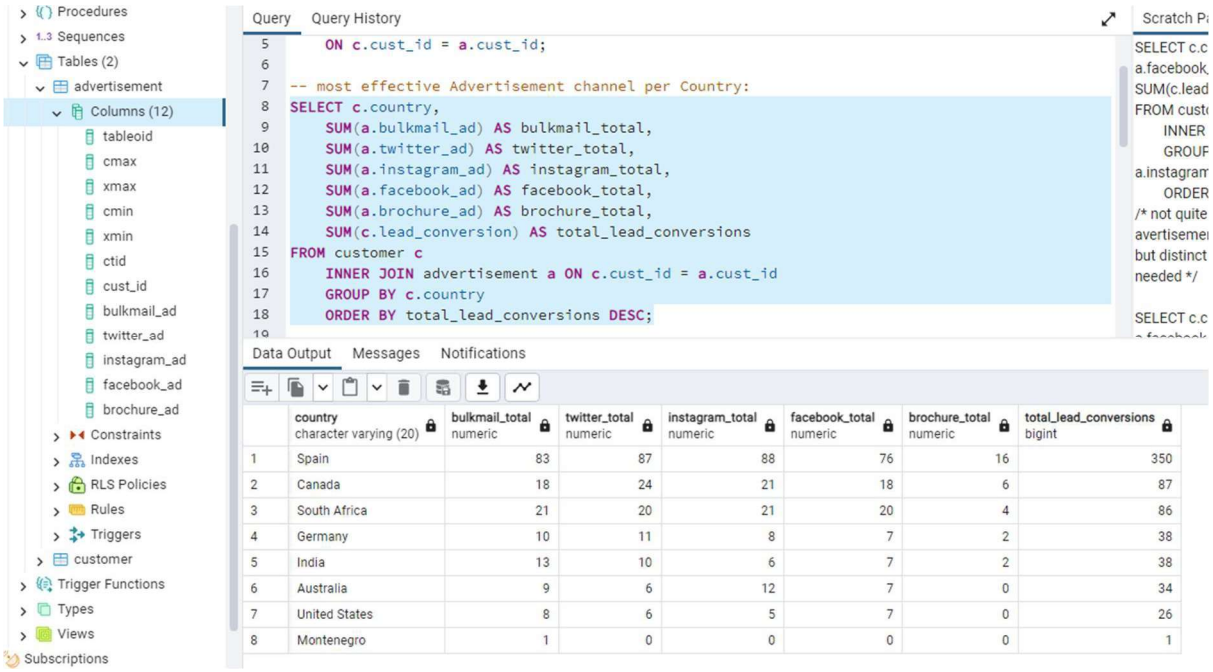
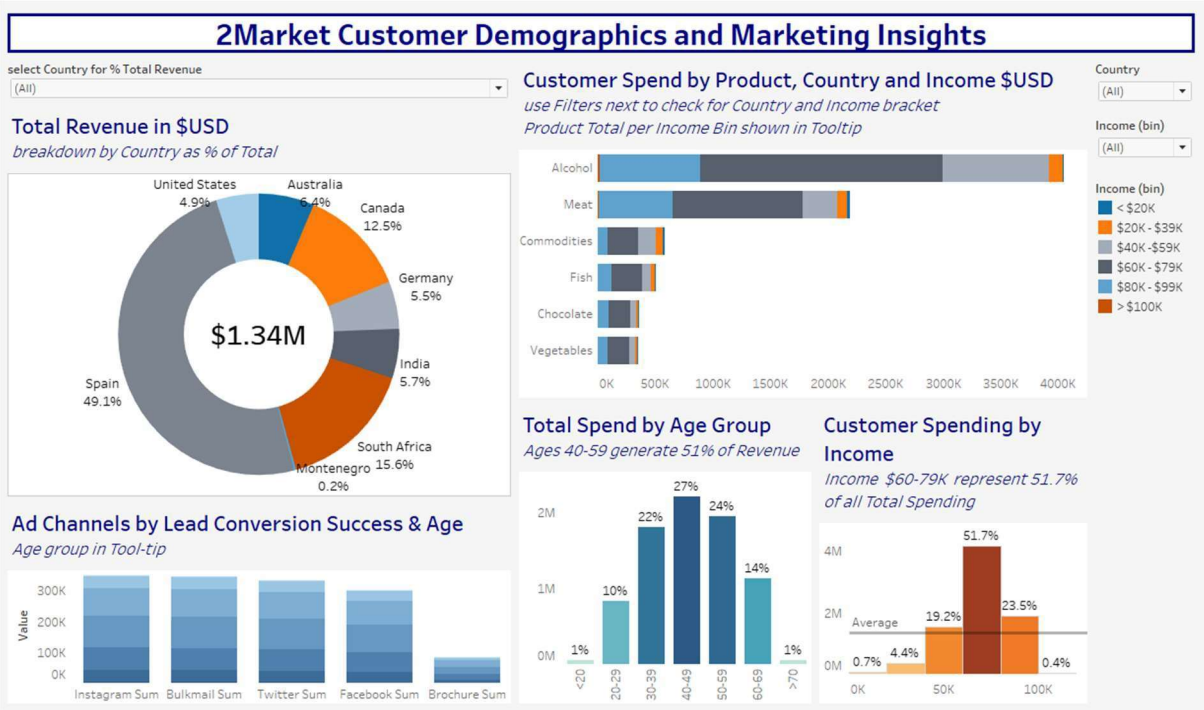


Tableau and PostgreSQL provided initial and in-depth insights. Facebook emerged as a consistently dominant advertising channel across countries. Instagram displayed exceptional effectiveness for specific product categories in Canada, Australia, and South Africa. Twitter exhibited niche impact for Alcohol and Chocolate products in Germany. Fish products consistently performed well across all channels in India. Montenegro, unfortunately, faced data limitations, urging the need for additional data collection or targeted market research.

Top Insights and Recommendations:

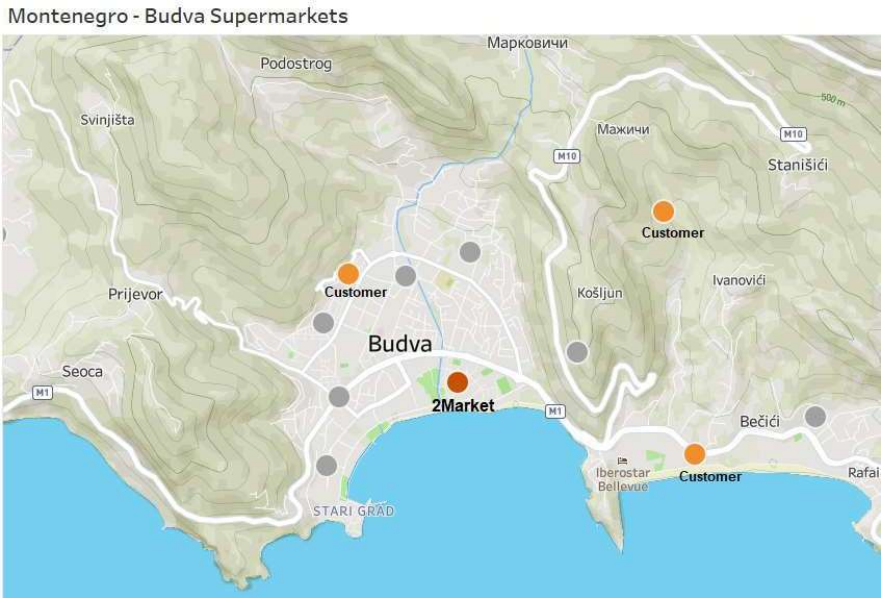
1. **Facebook Dominance:** Allocate substantial resources to Facebook advertising globally for increased reach and impact.
2. **Instagram Success:** Focus on Instagram for specific product categories in Canada, Australia, and South Africa.
3. **Twitter's Niche Impact:** Prioritise Twitter advertising for Alcohol and Chocolate products in Germany.
4. **Diverse Channel Impact:** Maintain a balanced approach across all channels for advertising Fish products in India.
5. **Montenegro Data Limitations:** Collect additional data or conduct targeted market research for more informed decision-making.
6. **Regional Variations:** Tailor advertising strategies based on regional preferences and product types for optimised impact.
7. **Consistent Facebook and Instagram Performance:** Continue leveraging Facebook and Instagram as staple channels in advertising strategies.
8. **Diversity in Product-Channel Dynamics:** Customise advertising strategies based on the specific dynamics of each product category.

Tableau Final Dashboard and Patterns: The final Tableau dashboard revealed a correlation between spend and discount deals, suggesting a trend that requires sales and marketing attention for appropriate discount selections.



Montenegro Challenge: Insufficient data from Montenegro highlighted the need for comprehensive insights. A geographic visualisation depicted hypothetical customer locations and the current competitive landscape, emphasising the importance of data-driven decisions in this region.

*Figure showing 2Market location in Budva – Montenegro. Orange dots marking Customer locations. Competitors in grey. [** Note: data consists of hypothetical Customer locations, plus actual supermarket data]*



Future Recommendations:

1. **Product Bundling:** Create bundled discounts for popular products like alcohol and meat to encourage higher spending.
2. **Loyalty Programs:** Introduce loyalty programs offering additional benefits for purchasing popular products.
3. **Cross-Selling Opportunities:** Explore cross-selling opportunities by suggesting complementary items to customers buying alcohol or meat.
4. **Montenegro Insights:** Obtain more comprehensive insights through targeted surveys, collaboration with local partners, and a deeper exploration of regional events, economic conditions, and cultural aspects.
5. **Spain Market Expansion:** Consider expanding marketing efforts in Spain, tailoring promotions based on local preferences, and exploring partnerships with local influencers or businesses to enhance brand visibility across social media.

This analysis serves as a strategic foundation for 2Market, paving the way for data-driven decisions and optimised advertising strategies in a dynamic post-financial crisis landscape.

In conclusion, effective advertising strategies for 2Market demand a nuanced and adaptable approach. Leveraging the strengths of each channel, considering regional variations, and continuous monitoring will enhance overall advertising effectiveness and contribute to improved sales performance.

Appendix

Column Headers renamed

Initial Column header	Replaced with
ID	cust_id
Year_Birth	birth_year
AGE_Customer	cust_age
Education	education
Marital_Status	marital_status
Income	income
Kidhome	nr_children
Teenhome	nr_teenagers
Dt_Customer	registration_date
Recency	purchase_recency
AmtLiq	alcohol_spend
AmtVege	vegetables_spend
AmtNonVeg	meat_spend
AmtPes	fish_spend
AmtChocolates	chocolate_spend
AmtComm	commodities_spend
NumDeals	nr_discount_deals
NumWebBuy	nr_website_buy
NumWalkinPur	nr_instore_buy
NumVisits	monthly_website_visits
Response	campaign_response
Complain	cust_complaint
Country_key	country_key
Country	country
Count_success	lead_conversion

SQL Syntax

```
-- create Customer table
CREATE TABLE Customer (
    cust_id INTEGER PRIMARY KEY,
    birth_year INTEGER,
    cust_age INTEGER,
    education VARCHAR (10),
    marital_status VARCHAR (10),
    income NUMERIC,
    nr_children INTEGER,
    nr_teenagers INTEGER,
    registration_date DATE,
    purchase_recency INTEGER,
    alcohol_spend NUMERIC,
    vegetables_spend NUMERIC,
```

```
    meat_spend NUMERIC,
    fish_spend NUMERIC,
    chocolate_spend NUMERIC,
    commodities_spend NUMERIC,
    nr_discount_deals INTEGER,
    nr_website_buy INTEGER,
    nr_instore_buy INTEGER,
    monthly_website_visits INTEGER,
    campaign_response NUMERIC,
    cust_complaint NUMERIC,
    country_key VARCHAR (3),
    country VARCHAR (20),
    lead_conversion INTEGER
);


-- create ADs table
CREATE TABLE Advertisement (
        cust_id INTEGER PRIMARY KEY,
        bulkmail_ad NUMERIC,
        twitter_ad NUMERIC,
        instagram_ad NUMERIC,
        facebook_ad NUMERIC,
        brochure_ad NUMERIC
);

-- Joining the Tables
SELECT * FROM customer c
        INNER JOIN advertisement a
        ON c.cust_id = a.cust_id;

-- most effective Advertisement channel per Country
SELECT c.country,
    SUM(a.bulkmail_ad) AS bulkmail_total,
    SUM(a.twitter_ad) AS twitter_total,
    SUM(a.instagram_ad) AS instagram_total,
    SUM(a.facebook_ad) AS facebook_total,
    SUM(a.brochure_ad) AS brochure_total,
    SUM(c.lead_conversion) AS total_lead_conversions
FROM customer c
        INNER JOIN advertisement a ON c.cust_id = a.cust_id
        GROUP BY c.country
        ORDER BY total_lead_conversions DESC;

-- most effective Advertisement channel based on Marital Status
SELECT c.marital_status,
    SUM(a.bulkmail_ad) AS bulkmail_total,
    SUM(a.twitter_ad) AS twitter_total,
    SUM(a.instagram_ad) AS instagram_total,
    SUM(a.facebook_ad) AS facebook_total,
    SUM(a.brochure_ad) AS brochure_total,
    SUM(c.lead_conversion) AS total_lead_conversions
FROM customer c
        INNER JOIN advertisement a ON c.cust_id = a.cust_id
        GROUP BY c.marital_status
```

```
        ORDER BY total_lead_conversions DESC;

SELECT
  c.marital_status,
  a.social_media_platform,
  SUM(a.lead_conversions) AS total_lead_conversions
FROM customer_data c
INNER JOIN advertising_data a ON c.customer_id = a.customer_id
GROUP BY c.marital_status, a.social_media_platform
ORDER BY c.marital_status, total_lead_conversions DESC;


-- total spend per Country
SELECT c.country,
  SUM(c.alcohol_spend + c.chocolate_spend + c.fish_spend +
        c.meat_spend + c.vegetables_spend) AS total_spend
FROM customer c
        GROUP BY c.country
        ORDER BY total_spend DESC;

-- total spend per Product per Country
SELECT c.country,
  SUM(c.alcohol_spend) AS total_alcohol_spend,
SUM(c.chocolate_spend) AS total_chocolate_spend,
SUM(c.commodities_spend) AS total_commodities_spend,
SUM(c.fish_spend) AS total_fish_spend,
SUM(c.meat_spend) AS total_meat_spend,
SUM(c.vegetables_spend) AS total_vegetables_spend
FROM customer c
GROUP BY c.country
ORDER BY total_alcohol_spend DESC;

-- total spend per Product by Marital
SELECT c.marital_status,
  SUM(c.alcohol_spend) AS total_alcohol_spend,
SUM(c.chocolate_spend) AS total_chocolate_spend,
SUM(c.commodities_spend) AS total_commodities_spend,
SUM(c.fish_spend) AS total_fish_spend,
SUM(c.meat_spend) AS total_meat_spend,
SUM(c.vegetables_spend) AS total_vegetables_spend
FROM customer c
GROUP BY c.marital_status
ORDER BY total_alcohol_spend DESC;

-- total spend per Product by Children and Teens
SELECT c.nr_children, c.nr_teenagers,
SUM(c.alcohol_spend) AS total_alcohol_spend,
SUM(c.chocolate_spend) AS total_chocolate_spend,
  SUM(c.commodities_spend) AS total_commodities_spend,
SUM(c.fish_spend) AS total_fish_spend,
SUM(c.meat_spend) AS total_meat_spend,
SUM(c.vegetables_spend) AS total_vegetables_spend
FROM customer c
GROUP BY c.nr_children, c.nr_teenagers
ORDER BY total_alcohol_spend DESC;
```

```
/* Which advertisement channels seems to be the most effective per country? */
SELECT c.country,
  'Bulkmail' AS advertising_channel,
SUM(a.bulkmail_ad) AS total_ad_engagement,
SUM(c.lead_conversion) AS total_lead_conversions
FROM customer c
        INNER JOIN advertisement a ON c.cust_id = a.cust_id
        GROUP BY c.country

UNION

SELECT c.country,
  'Twitter' AS advertising_channel,
  SUM(a.twitter_ad) AS total_ad_engagement,
  SUM(c.lead_conversion) AS total_lead_conversions
FROM customer c
        INNER JOIN advertisement a ON c.cust_id = a.cust_id
        GROUP BY c.country

UNION

SELECT c.country,
  'Instagram' AS advertising_channel,
SUM(a.instagram_ad) AS total_ad_engagement,
SUM(c.lead_conversion) AS total_lead_conversions
FROM customer c
        INNER JOIN advertisement a ON c.cust_id = a.cust_id
        GROUP BY c.country

UNION

SELECT c.country,
  'Facebook' AS advertising_channel,
SUM(a.facebook_ad) AS total_ad_engagement,
SUM(c.lead_conversion) AS total_lead_conversions
FROM customer c
        INNER JOIN advertisement a ON c.cust_id = a.cust_id
        GROUP BY c.country

UNION

SELECT c.country,
  'Brochure' AS advertising_channel,
SUM(a.brochure_ad) AS total_ad_engagement,
SUM(c.lead_conversion) AS total_lead_conversions
FROM customer c
        INNER JOIN advertisement a ON c.cust_id = a.cust_id
        GROUP BY c.country
        ORDER BY country, total_ad_engagement DESC,
        advertising_channel,
total_lead_conversions DESC;

/* Amount spent per Product, per Country - by Advertisement channel.*/
```

```
SELECT * FROM products;
SELECT * FROM advertisement;

ALTER TABLE a.advertisement ADD COLUMN
country VARCHAR (20);
UPDATE advertisement a SET country = c.country
FROM customer c WHERE a.cust_id = c.cust_id;

-- returns NULL in Country for several entries
-- confirmed: these were the outliers from before
DELETE FROM advertisement WHERE country ISNULL;

SELECT country, products, Bulkmail, Twitter, Instagram, Facebook, Brochure
FROM (
SELECT a.country, 'alcohol' AS products,
	SUM(a.bulkmail_ad) AS Bulkmail,
	SUM(a.twitter_ad) AS Twitter,
	SUM(a.instagram_ad) AS Instagram,
	SUM(a.facebook_ad) AS Facebook,
	SUM(a.brochure_ad) AS Brochure
FROM public.advertisement a
INNER JOIN products p ON a.cust_id = p.cust_id
WHERE p.alcohol_spend > 0
GROUP BY a.country, p.alcohol_spend

UNION

SELECT a.country, 'chocolate' AS products,
	SUM(a.bulkmail_ad) AS Bulkmail,
	SUM(a.twitter_ad) AS Twitter,
	SUM(a.instagram_ad) AS Instagram,
	SUM(a.facebook_ad) AS Facebook,
	SUM(a.brochure_ad) AS Brochure
FROM public.advertisement a
INNER JOIN products p ON a.cust_id = p.cust_id
WHERE p.chocolate_spend > 0
GROUP BY a.country, p.chocolate_spend

UNION

SELECT a.country, 'commodities' AS products,
	SUM(a.bulkmail_ad) AS Bulkmail,
	SUM(a.twitter_ad) AS Twitter,
	SUM(a.instagram_ad) AS Instagram,
	SUM(a.facebook_ad) AS Facebook,
	SUM(a.brochure_ad) AS Brochure
FROM public.advertisement a
INNER JOIN products p ON a.cust_id = p.cust_id
WHERE p.commodities_spend > 0
GROUP BY a.country, p.commodities_spend

UNION

SELECT a.country, 'fish' AS products,
	SUM(a.bulkmail_ad) AS Bulkmail,
```

```
        SUM(a.twitter_ad) AS Twitter,
        SUM(a.instagram_ad) AS Instagram,
        SUM(a.facebook_ad) AS Facebook,
        SUM(a.brochure_ad) AS Brochure
FROM public.advertisement a
INNER JOIN products p ON a.cust_id = p.cust_id
WHERE p.fish_spend > 0
GROUP BY a.country, p.fish_spend

UNION

SELECT a.country, 'meat' AS products,
        SUM(a.bulkmail_ad) AS Bulkmail,
        SUM(a.twitter_ad) AS Twitter,
        SUM(a.instagram_ad) AS Instagram,
        SUM(a.facebook_ad) AS Facebook,
        SUM(a.brochure_ad) AS Brochure
FROM public.advertisement a
INNER JOIN products p ON a.cust_id = p.cust_id
WHERE p.meat_spend > 0
GROUP BY a.country, p.meat_spend

UNION

SELECT a.country, 'vegetables' AS products,
        SUM(a.bulkmail_ad) AS Bulkmail,
        SUM(a.twitter_ad) AS Twitter,
        SUM(a.instagram_ad) AS Instagram,
        SUM(a.facebook_ad) AS Facebook,
        SUM(a.brochure_ad) AS Brochure
FROM public.advertisement a
INNER JOIN products p ON a.cust_id = p.cust_id
WHERE p.vegetables_spend > 0
GROUP BY a.country, p.vegetables_spend
        ) AS subquery
ORDER BY country, products;
```