



# Data Analytics capstone project notes guide

## Project context:

- **Scenario:** Imagine yourself working at a marketing agency whose client is a telecommunications company who is outsourcing their reporting needs to your organization. Your client is a global organization who worked with multiple agencies to activate media across various channels. Your organization is responsible for their campaign reporting.
- **Task:** Your client needs your expertise to maximize their future strategies by analyzing their reporting for Location, Audience, Creative and Inventory insights and recommendations
- **Action:** Identify KPI based on Location, Audience, Creative and Inventory. Benchmark performance relative to client goals. Dive deeper into the data by identifying insights related to your high level findings
- **Result:** Create a one-sheeter, along with 2-3 slides with visualizations of your key findings and recommendations

## Client's objectives:

Your client has a live campaign and wants to ensure its optimal effectiveness. To this end, they've tasked you with:

1. **KPI Insights:** KPI (Key Performance Indicators) are metrics that measure the success of specific activities. Your client wants insights into how these metrics are performing. This could include metrics like click-through rates, conversion rates, impressions, etc.
2. **Data Visualization:** Visual representation of data, such as graphs, charts, or maps. It helps in understanding complex data sets and identifying patterns or trends.
3. **Pivot Tables:** A pivot table is a data summarization tool used in spreadsheet programs. It allows you to reorganize and summarize selected columns and rows of data to obtain a desired report.
4. **Live Dashboard:** A real-time interface that displays a visual summary of the campaign's performance. It should be updated in real time or near-real time to reflect the current status of the campaign.



5. **Recommendations + Next Steps:** Based on the analysis of the campaign data, provide actionable suggestions to improve performance, along with a roadmap or plan for implementing those recommendations.

### Your Workflow:

1. **Prepare:** This involves cleaning and organizing the raw data to make it usable. This might include removing duplicates, handling missing values, and ensuring data consistency.
2. **Process:** Transforming the prepared data to a format or structure suitable for analysis. This could involve aggregating data, segmenting it based on certain criteria, or applying statistical methods.
3. **Analyze:** Dive deep into the data to extract meaningful insights. This step involves using statistical tools, machine learning models, or other analytical methods to derive conclusions from the data.
4. **Share:** After analysis, it's time to communicate the findings. This involves creating reports, presentations, and sharing the live dashboard. The insights and recommendations derived from the analysis should be presented in a clear and concise manner to the client.

### End Goal:

To retain your client, you need to provide high-quality service by transforming raw data into actionable insights and clear recommendations. This process of turning raw data into meaningful information and insights will help the client maximize the effectiveness of their campaign.

**In summary**, you're being tasked with taking raw data from the client's campaign, processing and analyzing it, and then presenting your insights in a visually appealing and understandable format, along with actionable recommendations for improvement.



## Measuring effectiveness: Campaign Goals and Performance Metrics

Goal type	Primary goal	Secondary goal
Audience	\$250 CPA	0.015% CTR
Desktop conversion	\$350 CPA	0.015% CTR
Smartphone conversion	\$400 CPA	0.015% CTR
Viewability	60%	\$80 vCPM

### Table breakdown:

1. **Goal Type:** This column describes the type of audience or inventory being targeted.
  - **Audience:** Audience targeting in advertising refers to the practice of identifying and reaching specific groups of people who are likely to be interested in a product, service, or message.
  - **Desktop/Smartphone Conversion:** A conversion occurs when a visitor to your website completes a desired goal, such as filling out a form or making a purchase.
  - **Viewability:** Viewability in advertising refers to the measurement of whether or not an ad has the opportunity to be seen by a user. It is an important metric for advertisers because it indicates the likelihood that their ads are actually being viewed by their target audience.
2. **Primary Goal:** The main objective or KPI for each goal type.
  - **CPA ( Cost Per Acquisition):** In the case of cost per action, the company pays a fee when the ad results in an action taken by a customer. The action could be anything from subscribing to a newsletter to downloading an eBook. Actions generally refer to when consumers click on a particular button that takes them to a separate page or acts as a download link.
  - **Viewability:** A viewable impression is a standard measure of ad viewability defined by the International Advertising Bureau (IAB) to be an ad, which



appears at least 50% on screen for more than one second. Viewable impressions are the metric that advertisers use to quantify the percentage of ads that are actually viewed by real people.

3. **Secondary Goal:** Another important metric or KPI for each goal type.
  - **CTR:** Click-Through Rate. The percentage of ad views that result in a click.
  - **vCPM:** Cost Per Thousand Viewable Impressions. The price of 1,000 advertisement impressions on web pages that are viewable.

## Objectives + Goals:

Each pod member has a unique role and will analyze data sets based on their respective area of focus. Goals are consistent to align across marketing initiatives, **each pod member should analyze performance relative to their role focus area - while actively** cross-collaborating and communicating among and across team members.

**Location Strategy:** Analyze location performance relative to campaign KPIs.

**Identify which locations were the most **cost effective****

- Performance metrics (CPA, CVR, CPC, CPM)
- Branding metrics (CTR, viewability, viewable impressions, etc.)

**What are the top/bottom performing **STATES?****

- Create an additional column to group and sort by State
- Provide recommendations to maximize or minimize location based targeting through state level targeting

**Explore **additional insights** beyond location to inform your team**

- What are top audiences, websites/exchanges, devices, creatives, viewability - across top/bottom performing locations?



**Audience Goals:** Analyze audience segment's performance relative to campaign KPIs.

**Identify top 10 performing audiences and create at least three audience personas**

- Provide evidence to support your strategy
- Consider exploring outside sources to justify your findings (news, trends, government regulations, etc. )

**Create recommendations for testing strategies**

- Recommend audiences who are worth testing in future strategies

**Clean up your data sets**

- Review raw data and make formatting adjustments  
EX. replace " &gt;" with " >"

**Identify which audiences were the most **cost effective****

- Performance metrics (CPA, CVR, CPC, CPM)
- Branding metrics (CTR, viewability, viewable impressions, etc.)

**Explore **additional insights** beyond Audience to inform your team**

- What are top audiences, websites/exchanges, devices, creatives, viewability - across top/bottom performing audiences?

**Creative Strategy:** Analyze creative performance relative to campaign KPIs.

- **Identify which creative sizes are the highest and lowest performers**
  - Performance metrics (CPA, CVR, CPC, CPM)
  - Branding metrics (CTR, viewability, viewable impressions, etc.)
- **Inform your client on what sizes are the highest and lowest performers by**
  - Device
  - Location
  - Messaging



**Inventory & Viewability:** Analyze inventory and viewability performance relative to campaign KPIs.

- **Identify which inventory sources (exchanges ) yield the greatest performance in terms of:**
  - Performance metrics (CPA, CVR, CPC, CPM)
  - Branding metrics (CTR, viewability, viewable impressions, etc.)
  - Impressions (Viewable vs measurable )
    - What is the difference between using measurable vs general impressions?
- **Identify at least five different publishers (websites) whom your client should contract for a direct buy**
  - Why these publishers?
  - What rate should you pay for a direct buy?

**Identify 5 Publishers:** Based on the inventory performance and viewability metrics, recommend at least 5 publishers that offer valuable ad placements.

- **Rationale for Selection:** Provide reasons why these publishers are recommended. This could be due to high viewability rates, user engagement, audience alignment, etc.

**Direct Buy Rate Recommendations:** Suggest the rate or price the client should pay these publishers for direct ad placements. This rate should ensure good ROI while also being competitive and fair for the value provided by the publisher.

## Metrics:

1. **CPA (Cost Per Acquisition):** This metric indicates how much it costs to acquire one customer or lead.  $CPA = \text{Total Ad Spend} / \text{Total Acquisitions}$ 
  - For instance, if you spent \$1000 on ads and acquired 10 customers, the CPA would be \$100.
2. **CTR (Click-Through Rate):** Represents the percentage of people who clicked on the ad out of the total who viewed it.
  - $CTR = (\text{Number of Clicks} / \text{Number of Impressions}) \times 100\%$
3. **CPC (Cost Per Click):** This is the cost paid for each click on the ad.
  - $CPC = \text{Total Ad Spend} / \text{Total Clicks}$



4. **CPM (Cost Per Mille/Thousand):** Represents the cost of 1,000 advertisement impressions.
  - $CPM = \text{Total Ad Spend} / \text{Total Impressions in Thousands}$
5. **CVR (Conversion Rate):** Indicates the percentage of users who take a desired action (like making a purchase or signing up).
  - $CVR = (\text{Number of Conversions} / \text{Total Visitors}) \times 100\%$
6. **vCPM (Viewable Cost per Mille/Thousand):** Measures how frequently an ad is seen by users, as opposed to the number of times it's placed by the seller.
  - $vCPM = 1000 \times \text{Cost} / \text{Viewable Impressions}$

These formulas are the mathematical representations of the metrics you're looking to calculate, and they are fundamental when evaluating advertising campaign performance data.

## Functions & Deliverables:

### 1. Consistent Naming Conventions:

- **Purpose:** This is foundational for any data analytics project. It ensures that all components of your dataset are labeled in a consistent and standardized manner. Alignment on naming conventions may negatively impact cross-collaboration when working with various teams who do not align on clear and consistent naming conventions. EX: The simple action of a user signing up for a newsletter. You could implement the event as "Sign up," "Signup," or "User Signed Up."
- **Benefits:**
  - Integration: When integrating data from various sources, consistent naming helps avoid confusion and errors.
  - Readability: It makes datasets more user-friendly, especially for team members who might not be deeply familiar with the data.
  - Automation: Consistency aids in automation processes, as scripts or tools can rely on standardized names.

### 2. Pivot Tables:

- **Purpose:** Pivot tables are used to summarize, analyze, and present data in a more digestible format. They allow you to take large datasets and transform them into a summarized version based on your chosen parameters.
- **Benefits:**



- **Data Aggregation:** You can quickly aggregate data to understand trends and patterns.
- **Flexibility:** They allow for dynamic data exploration by dragging and dropping fields.
- **Simplification:** Helps in converting complex data views into simpler tables that highlight the most important aspects.

### 3. **Vlookups:**

- **Purpose:** The VLOOKUP function is used to search for a value in the first column of a table or range. Once found, it returns a value in the same row from a specified column.
- **Benefits:**
  - **Data Matching:** Useful for comparing two lists or tables and finding matching or corresponding data.
  - **Data Enrichment:** If you have supplemental data in a separate table, VLOOKUP can help you integrate this additional data into your main table.
  - **Error Detection:** May be used to identify discrepancies between datasets.

### 4. **Calculated Fields:**

- **Purpose:** These are custom fields in a table that you create. The value of this field is computed based on a formula that you define, which can use other fields in the table.
- **Benefits:**
  - **Custom Metrics:** Allows you to generate new metrics that aren't directly available in the raw data.
  - **Dynamic Calculation:** The values in calculated fields update automatically if the values of the dependent fields change.
  - **Flexibility:** Can be tailored to specific analytical needs, enhancing the depth of analysis.

### 5. **Dashboard:**

- **Purpose:** A dashboard is a visual tool that presents data in a graphical manner, such as charts, graphs, and tables. It's designed to offer a quick snapshot of key performance indicators and metrics.
- **Benefits:**





- Real-time Insights: Many dashboards update in real-time, allowing for up-to-the-minute insights.
- Interactivity: Users can often drill down into specific metrics, change time frames, or apply filters to view data from different angles.
- Decision Making: A well-designed dashboard provides clarity, making it easier for decision-makers to understand trends, challenges, and opportunities.

In summary, these functions and deliverables are essential tools in the data analytics process. They streamline data handling, enhance the clarity of insights, and facilitate more informed decision-making.

## Team Roles & Responsibilities

R&R expanded beyond Project Deck which mimics real world responsibilities as well as during the Capstone Project.

### 1. Spatial Analyst:

**Role:** A location analyst, also known as a geographic information systems (GIS) analyst or spatial analyst, is a professional who specializes in analyzing geographical data to provide insights and solutions related to location-based issues. Their primary role revolves around gathering, organizing, analyzing, and interpreting spatial data to support decision-making processes in various fields such as urban planning, environmental management, real estate, retail, logistics, and more.

#### Responsibilities:

- Takes notes during cross team meetings and collaborates across pod teams to exchange relevant details
- Specializes in analyzing geographical data to provide insights and solutions related to location-based issues
- Gathering, organizing, analyzing, and interpreting spatial data to support strategy recommendations
- Create maps and visualizations to represent spatial data effectively, including thematic maps, heat maps, and 3D models based on KPI
- Assess the demand for goods or services within specific geographic areas and provide recommendations for market expansion or targeting



## 2. Creative Analyst:

**Role:** Creative analysts may analyze the performance of marketing campaigns or initiatives to optimize effectiveness and ROI. This could involve A/B testing, multivariate analysis, or other experimentation methods to refine messaging, targeting, or creative elements.

### **Responsibilities:**

- Creative analyst work with data to uncover patterns, trends, and insights that inform decision-making based on creative performance by analyzing creative performance to inform strategies
- Gathering, organizing, and analyzing data to answer questions and identify key trends related to creative deliverables
- Create reports, presentations, infographics, or visualizations to communicate insights or findings effectively based on creative insights
- Support shaping brand messaging, positioning, and brand identity to resonate with target audiences based on creative insights and KPIs

## 3. Audience Analyst:

**Role:** An audience analyst is primarily responsible for gathering, analyzing, and interpreting data related to audience behavior, preferences, and engagement across various platforms or channels. Their insights are crucial for informing strategic decisions related to marketing, content creation, product development, and audience engagement.

### **Responsibilities:**

- Primarily responsible for gathering, analyzing, and interpreting data related to audience behavior, preferences, and engagement across various platforms or channels
- Audience analysts collaborate with various teams and departments, including marketing, advertising, content creation, product development, and sales, to align audience insights with business objectives and initiatives
- Segment audiences into distinct groups based on demographic factors (age, gender, location), psychographic characteristics (interests, attitudes, lifestyles), and behavioral attributes (purchase history, online activity)
- Create detailed profiles or personas of different audience segments to better understand their needs, preferences, and pain points based on KPIs

## 4. Inventory Analyst:

**Role:** An inventory analyst plays a critical role in optimizing inventory levels, improving inventory accuracy, and ensuring that inventory management processes support the company's overall goals and objectives.

**Responsibilities:**

- Conceptualizing, designing, and implementing data frameworks to build out Inventory reporting
- Responsible for managing and optimizing inventory supply and demand based on performance and KPI
- Identifies opportunities to optimize inventory supply by working closely with sales, marketing, and production teams ( within pod ) to understand demand drivers and incorporate demand forecasts into inventory planning processes
- Collaborate across departments to coordinate inventory optimization activities and address inventory-related issues effectively by gathering context from other teams

In the context of the capstone project, the collaboration between these roles is crucial for the holistic success of the advertising campaigns. Each role brings a specific expertise, and their combined insights help in making data-driven decisions.

## Additional Reference

### Market Efficiency

**Objective:** To understand which markets are most and least effective for the client's campaigns.

- **Top 10 Performing DMAs (Designated Market Areas):** Identify the top 10 areas where the campaign has performed the best.
- **Bottom 10 Performing DMAs:** Identify the 10 areas where the campaign has performed the least effectively.
- **Exclusion of Markets:** Based on performance metrics, recommend markets that should be excluded from future campaigns to save budget and increase overall efficiency.

### Device Types & Operating Systems:

**Objective:** To gauge the performance of various devices and their associated operating systems and inform future targeting.

- **Top and Bottom Devices:** Identify the top 10 and bottom 10 devices based on performance metrics.
- **Performance Analysis:** Investigate why certain devices are either performing well or underperforming. This could be due to device



capabilities, user demographics, or other factors.

- **Device Exclusion Recommendations:** Suggest devices that the client should possibly exclude from future campaigns based on their performance and the reasons for their underperformance.

### Device Performance in Different DMAs (locations):

**Objective:** To understand how devices perform specifically in the best and worst performing markets.

- **Top Devices in Top DMAs:** Identify the best-performing devices in the top 10 DMAs.
- **Top Devices in Bottom DMAs:** Identify the best-performing devices even in the least effective DMAs.

### Data Formatting & Naming:

**Objective:** Ensure clarity and consistency in the data, especially concerning tablets which might be grouped under mobile strategies.

- **Rename Strategies:** Review raw data and make formatting adjustments
  - i. EX. replace " »" with " >"

### Cost-Effectiveness of Devices and Locations:

**Objective:** To determine which devices and locations provide the best value for the money spent.

- **Gross Cost:** Includes the total cost of all media activity.

### Top Performing Audiences & Personas:

Understand which audiences resonate most with the campaigns and create representative profiles.

- **Top 10 Audiences:** Identify the best-performing audience segments.
- **Audience Personas:** Based on the top-performing audiences, create at least 3 detailed personas to represent these segments for better-targeted campaigns.



## Evidence & External Validation:

**Objective:** Ensure that the chosen strategies and audience segments are backed by solid evidence and external validation.

- This could involve referencing news articles, trends, or government regulations that align with the campaign's observations and findings.

## 2. Naming Conventions:

**Objective:** Ensure clarity and consistency in naming conventions for different audience segments.

## Creative Format Performance:

**Objective:** Understand which types of creative ads (like banner ads, video ads, interstitials, etc.) are most and least effective.

- **Top and Bottom Creative Formats:** Identify the best and worst-performing ad formats. This will help in optimizing ad spend by focusing on the most effective formats and possibly phasing out or improving the underperforming ones.

## Size Performance Analysis:

**Objective:** Determine which sizes of creatives are most effective, as the size can influence visibility and engagement.

- **Device:** Different devices (like mobiles, tablets, desktops) might have different optimal ad sizes.
- **Operating System:** Certain sizes might be more effective on, for example, iOS compared to Android, due to design standards or user behavior.
- **Location:** Ad sizes might perform differently in various geographical regions or DMAs (Designated Market Areas).
- **Strategy (line item):** Depending on the campaign strategy or specific campaign line items, certain ad sizes might be more effective.