**A picture containing text, tableware, plate, dishware

Description automatically generated**

**Career Challenge Guide**

Contents

[Welcome 2](#_Toc116483669)

[Your Career Challenge 3](#_Toc116483670)

[Challenge goals: 3](#_Toc116483671)

[Understanding the data 4](#_Toc116483672)

[Survey Data: 4](#_Toc116483673)

[Mosaic UK 7 Grand Index: 5](#_Toc116483674)

[Google Analytics Simulation: 5](#_Toc116483675)

[Different Approaches to take 6](#_Toc116483676)

[Approach A: Can you help the Data Strategy team? 6](#_Toc116483677)

[Approach 2: Can you help the Data Science team? 6](#_Toc116483678)

[Presenting your findings 6](#_Toc116483679)

# Welcome

Welcome to DigData live online career challenge. This document will help you understand the challenge and give you all information you need to know.

You should have received the following five documents before the live online session:

|  |  |
| --- | --- |
| Document | Description |
| Digdata Challenge Guide (this document) | This describes the assignment and the adjacent dataset and tells you how to get started on the challenge |
| Survey (xlsx) | The sample data you use to do the first part of the challenge |
| Mosaic UK 7 Grand Index (xlsx) | A document containing one of the segmentations from Experian. You can find key features for the different segments and percent separation by given criteria. |
| Google analytics simulation (xlsx) | A csv containing a few events to be observed in the second part of the challenge simulating the real data from Google Analytics. |
| Presentation (pdf) | Slides from Digdata presentation in case you want to review things we talked about |

Before you begin the challenge, make sure you understand the assignment and ask questions if necessary. Make sure you are provided with enough time, peace and quiet so you can focus and cover different sides of the given problem.

# Your Career Challenge

**How do you make an Electric Vehicle release a Success?**

*In this challenge you are BMW’s marketing agency preparing the launch of a new BMW Luxury Electric Vehicle (EV) priced at £70,000. BMW would like to attract potential buyers to come and have a test drive in the new EV. Use the data provided to recommend to BMW where they should focus their £100,000 marketing budget so as to attract as many test drivers as possible.*

*BMW have told you that from their previous experience potential buyers are most likely to live in urban areas, with access to driveways. They have also said that potential buyers are likely to be those who either already own a BMW or those who own a similar brand (e.g. Audi or Mercedes-Benz).*

# Challenge goals:

1. Use the questionnaire responses (Survey data) to determine the type of person who is most likely to be interested in test driving an EV. This will be your target audience
2. Using the Mosaic UK 7 Grand Index, find the “Group Type(s)” most similar to your targeted audience. What key characteristics do your Group Type(s) share and what areas are they most likely to be living in? Calculate how many prospects are likely to be arriving at BMW’s website and how much it will cost
3. Using the Google analytics data to identify 3 different buckets (groups).
   * 1. Non Prospects
     2. Prospects for re-targeting
     3. Prospects for a later campaign

# Understanding the data

## Survey Data:

Use the survey data to help determine what a future BMW customer might look like.

The raw data contains 7,250 rows (customers’ responses) and 13 columns (features).

The data is not real Experian data but is similar to what we use to help make business recommendations.

The dataset contains:

* Personal information, e.g. name or gender
* Driving related information, e.g. whether they have a driver’s license
* Potential buyer information , e.g. intention on buying, the price range they are aiming for, preferred make etc

An exhaustive list of columns is described below:

|  |  |
| --- | --- |
| Columns | Description |
| **first\_name** | Contains the first name of the responder |
| **last\_name** | Contains the last name of the responder |
| **gender** | Contains the gender of the responder |
| **age** | Contains the age of the responder |
| **town** | Contains the town of the responder |
| **income\_band** | Contains information about the responder's income |
| **preferred\_make** | Contains the responder’s preferred car make |
| **owns\_car** | Whether the responder owns a car or not |
| **has\_drivers\_license** | Whether the responder has a drivers license |
| **decision\_maker** | Whether the responder is the head of the household and can make decision on purchase or not |
| **intent\_on\_buying** | Whether the responder intends on buying or not |
| **has\_ongoing\_lease** | Whether the responder has an on-going lease or not |
| **target\_price\_range** | What price range the responder can afford |
| **has\_used\_electric\_vehicle** | Whether the responder has used electric vehicles or not |
| **has\_purchased\_from\_auto\_displays** | Whether the responder has purchased from auto displays or not |
| **requests\_test\_drive** | Whether the responder requests a test drive or not |
| **has\_driveway** | Whether the responder has a driveway or not |

## Mosaic UK 7 Grand Index:

This document contains multiple sheets representing different data that you could use for your analysis.

|  |  |
| --- | --- |
| Sheet | Description |
| Primary Content | There you can see general information about the segments – code, name, one-line description of the people segmented there, what percent of the UK population or household are distributed in this segment. |
| Key Features | Each group is represented by 6 key features for the people segmented in it |
| Grand Index – Means | Percentage distribution of the population in all the groups by variables related to the National Means |
| Dataset Pricing | Sample Pricing for all the datasets (variables). This is useful when you would like to get the information for a specific criterion. For example, if a company would like to receive the means distribution for “age > 18-25” they would need to pay £0.40 for 1,000 impressions (CPM). CTR stands for “click through rate”. So, if you served 1000 ad impressions and 100 people clicked on the ad, the CTR would be 10%. |
| Regional Distributions | shows the distribution of Mosaic groups for each region of the UK |

## Google Analytics Simulation:

Represents sample response data from BMW’s website traffic which is focused on specific events related to the research we are doing. The table below will help you understand what each column represents.

|  |  |
| --- | --- |
| Sessions | The number of times a user has visited a website |
| Average Session Duration | The average time a user spent on the website, per session |
| Bounce Rate | The percentage of sessions where a user only viewed one page, then left.  If a user had 1 session and a 100% bounce rate it would mean they only viewed one page.  if a user had 2 sessions and 50% bounce rate it would mean the user visited more than one page on 1 session, and only 1 page on the other session |
| Event 1 | User configured their EV |
| Event 2 | User visited the test drive sign up page |
| Event 3 | User visited the test drive sign up page, filled in their details and requested a test drive |
| Event # Conversion Rate | Percent of sessions which completed the event |

# Different Approaches to take

There are many ways to approach this task, depending on where your skillset lies. Below are two different approaches you could take:

## Approach A: Can you help the Data Strategy team?

Use the data to recommend what BMW should do. The team have suggested that you use the following steps to choose the most appropriate target group:

1. Find common characteristics between the people responding to the survey. Try to find anything common between the ones that are requesting a test drive.
2. Use Mosaic UK 7 Grand Index to identify which group/type is most similar to your targeted audiences and make a profile of a potential buyer.
3. Estimate the cost of an advertising campaign targeting your potential buyer profile. The “dataset pricing” is the cost for the information you would need to receive from a company such as Experian. Each characteristic’s data costs a varying amount of money. How many website viewers do you expect from your campaign?

## Approach 2: Can you help the Data Science team?

Use mathematical models and algorithms to direct business decisions. The team have suggested you use the following steps:

1. Identify a potential buyer profile (using the Survey data and Experian’s Mosaic segments)
2. Create an algorithm using the Google Analytics Data which will help to retarget BMW website users.
   1. Split the website visitors into 3 buckets depending on their activity.
      1. Low engagements, not prospects
      2. High engagement, prospects for re-target
      3. High engagement, prospects to be re-targeted at a later date
   2. Write a rule/algorithm that will automatically put the responses into the buckets using And, Or, Not statements

# Presenting your findings

Once you’ve come up with your analysis and ideas it will be time to present your results. Imagine your audience is a business leader who may not have the same background as you, so make sure to describe your recommendations in a clear and compelling way! You could do this in many ways, for example by using:

* Infographics and pictures
* Graphs and charts
* Code, text, or bullet points
* Video

**Submit all your work to** [**info@digdata.online**](mailto:info@digdata.online) **by Friday 21st October, and you will receive your virtual work experience certificate within 7 days**