



Technical Report

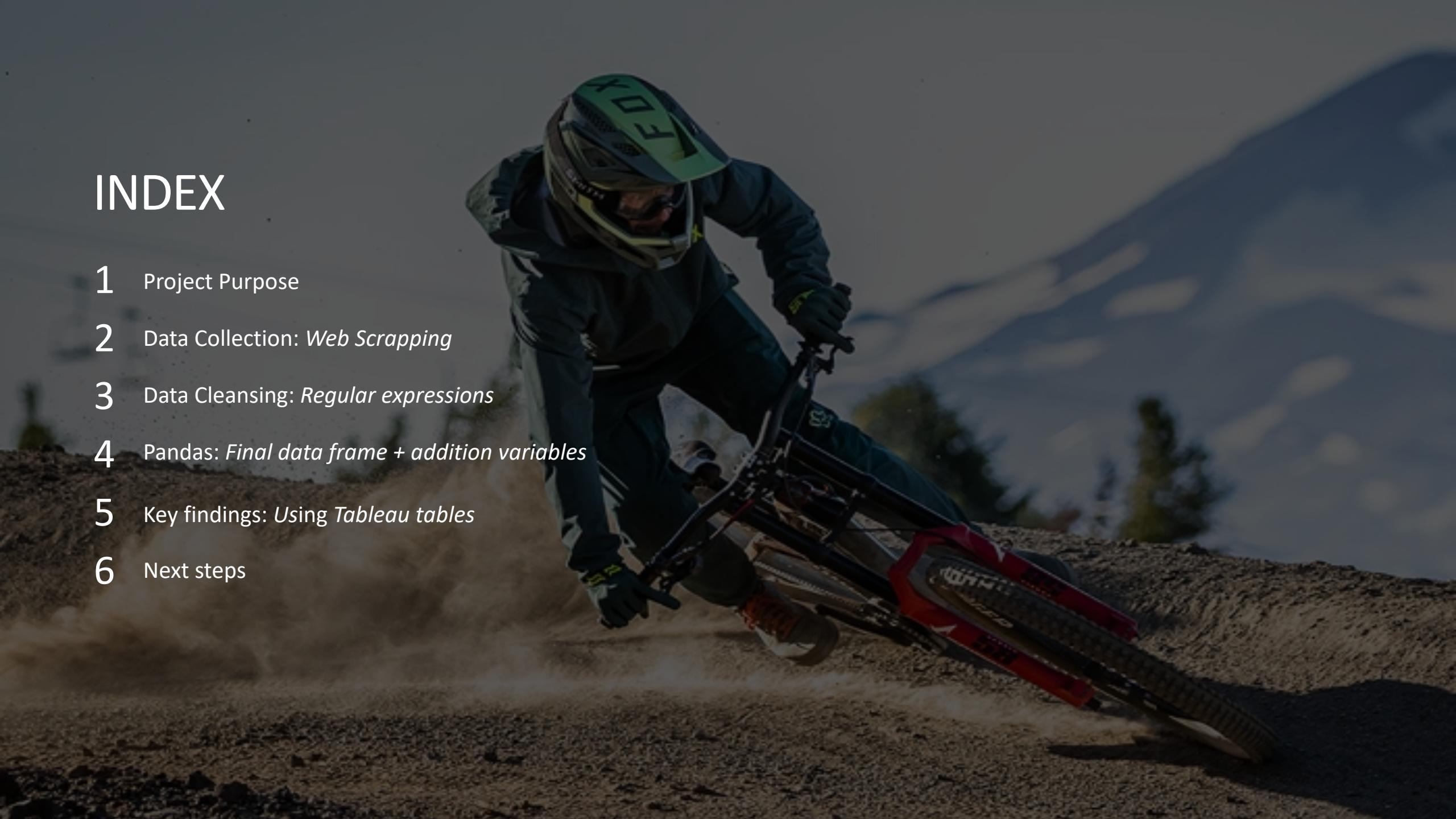
Mountain Bike Helmets Marketplace Study using Chain Reaction Cycles

11/23/2020

André Santa Clara



INDEX

- 
- 1 Project Purpose
 - 2 Data Collection: *Web Scrapping*
 - 3 Data Cleansing: *Regular expressions*
 - 4 Pandas: *Final data frame + addition variables*
 - 5 Key findings: *Using Tableau tables*
 - 6 Next steps

INDEX

- 1** Project Purpose
- 2** Data Collection: *Web Scrapping*
- 3** Data Cleansing: *Regular expressions*
- 4** Pandas: *Final data frame + addition variables*
- 5** Key findings: *Using Tableau tables*
- 6** Next steps



This project takes all information on MTB helmets from Chain Reaction Cycle website for the purpose of getting a better understanding of how Fox Racing helmets are being represented in their e-commerce.



MTB Helmets

In this study scope focuses on gathering **MTB Helmet** data from an online marketplace.



Data used

The market place chosen was **Chain Reaction Cycles**, one of the largest market places in the MTB world. The data extracted comes from the **MTB Helmet category** as well as the subcategories within the category



Data Collection

Web-scraping was conducted on the Chain Reaction Cycles website to take information regarding MTB Helmets



Analysis Variables

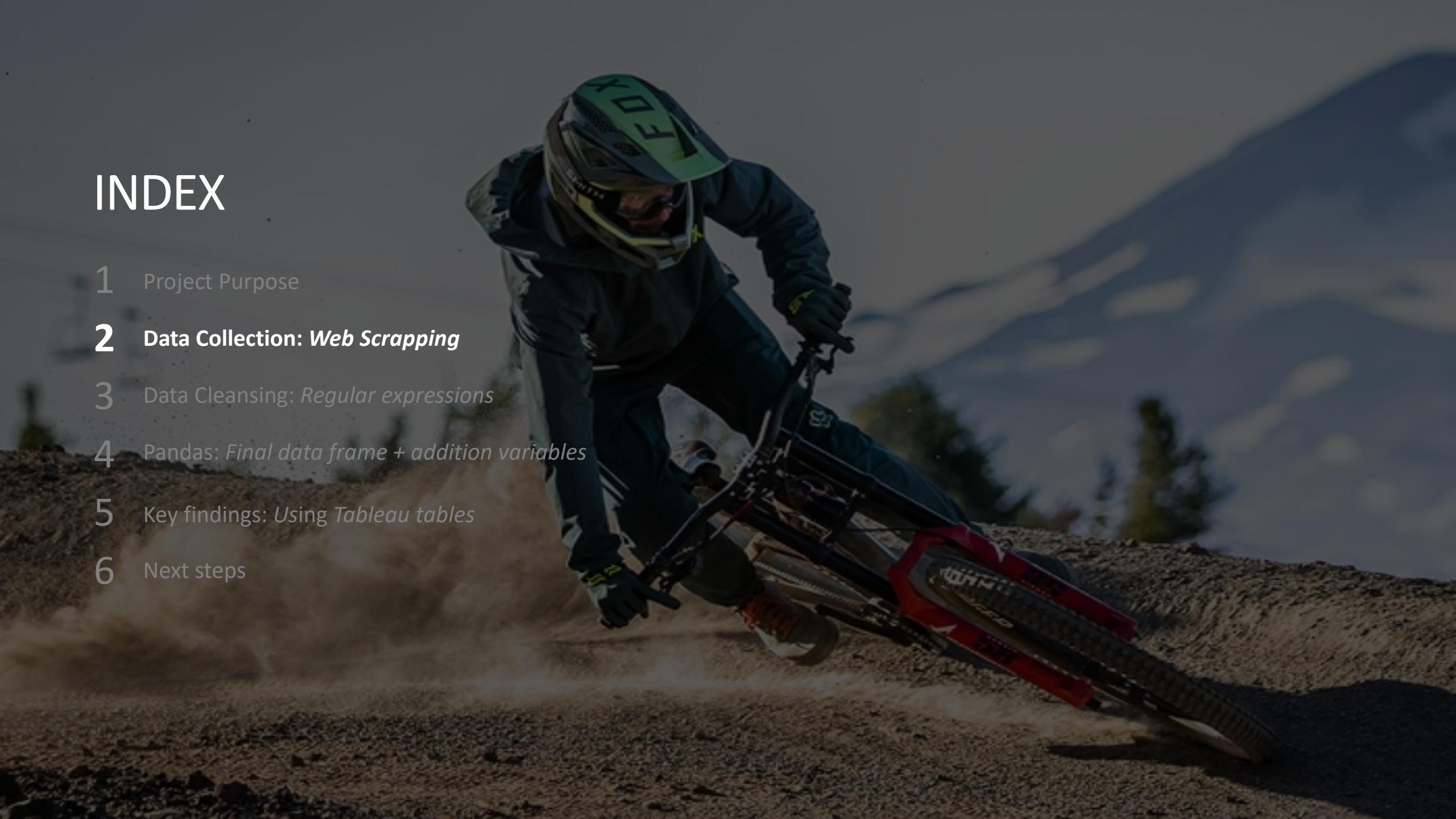
The following variable were gathered from all MTB helmets sold on the CRC website: product description, sub-category, brand, price & visibility score (product positioning on website). Most variable information was extracted by the product descriptions



Data Treatment

Web scrapping and data cleansing were done using python code. Graphs were produced using Tablaeu

INDEX

- 
- 1 Project Purpose
 - 2 Data Collection: *Web Scrapping***
 - 3 Data Cleansing: *Regular expressions*
 - 4 Pandas: *Final data frame + addition variables*
 - 5 Key findings: *Using Tableau tables*
 - 6 Next steps

Web-scraping was used to extract Mountain Bike helmet data from Chain Reaction Cycles (CRC). Five total websites were scrapped: the main MTB helmet page (with all the helmet products) and the four individual MTB helmet sub-category pages (for the purpose of adding the sub-category variable)

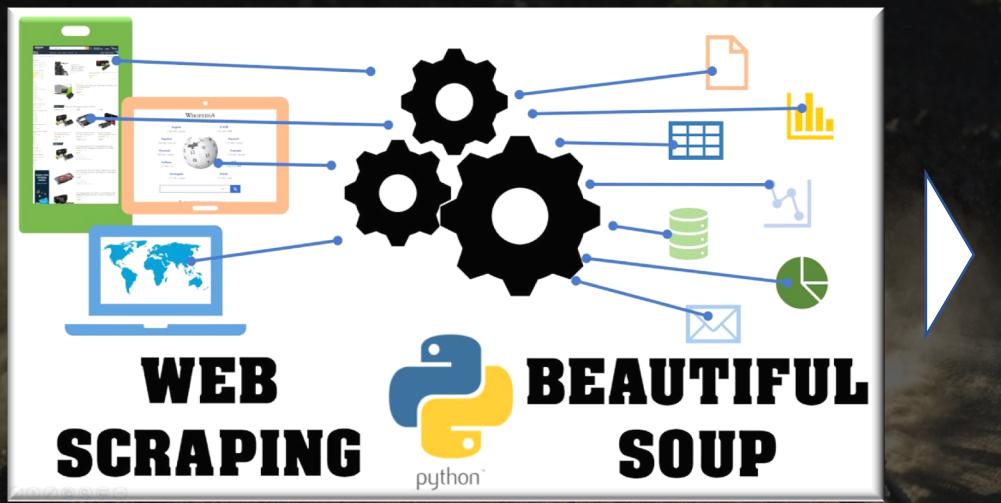
General search page

The screenshot shows the main navigation bar with links for Sign In, My Account, Help, Wishlist, and currency selection (GBP). Below the navigation is a search bar and a 'SEARCH' button. A 'BLACK FRIDAY' banner is prominently displayed. The main content area is titled 'Helmets' and shows a grid of three MTB helmets: 'SixSixOne Reset Helmet', 'IXS Trail EVO Helmet Exclusive 2020', and 'SixSixOne Reset MIPS Helmet'. Each item has a price range, review count, and a 'SAVE UP TO 50%' badge.

Sub-category pages: MTB helmets, Full-Face, Kids, dirt-jump

The screenshots show two sub-category pages: 'Full Face Helmets' and 'MTB Helmets'. Both pages feature a 'BLACK FRIDAY' banner and a 'WEEK 3' promotional section. The 'Full Face Helmets' page displays three helmets: 'SixSixOne Reset Helmet', 'SixSixOne Reset MIPS Helmet', and 'Brand-X DH11 Full Face Helmet'. The 'MTB Helmets' page displays three helmets: 'IXS Trail EVO Helmet Exclusive 2020', 'Leatt MTB 2.0 Helmet 2021', and 'Fox Racing Dropframe Pro MTB Helmet AW20'. Each product listing includes a price, review count, and a 'SAVE' badge.

For web-scraping the Beautiful Soup library was used. The first step web-scraping is understanding the information we want to extract: product description, sales price and original price. Secondly, it was important to inspect the websites html to successfully extract the information we were looking for.



Main MTB helmets page. with No sub-categories

```
In [2]: baseUrl = "https://www.chainreactioncycles.com/mtb/helmets?f=2072&page={}"  
nPages = 6  
  
In [3]: name = []  
price = []  
oldPrice = []  
  
for i in range(1,nPages+1):  
    page = BeautifulSoup(requests.get(baseUrl.format(i)).content, "html.parser")  
  
    for item_selector in page.select(".products_details.product_details_plp"):  
        name.append(item_selector.h2.text.strip())  
        try: price.append(item_selector.select_one(".fromamt").text.strip())  
        except: price.append("undefined")  
  
        try: oldPrice.append(item_selector.select_one(".rrpamount").text.strip())  
        except: oldPrice.append("Undefined")
```

Example

New Learning / Tip 🔎

- Circled in blue is the methodology used to parse through website that contained 6 different pages. I found that the different pages url simply changed by the number at the end of the url: "pages=1". Thus, using "{}" as a place holder in the baseURL we are able to use a for loop to parse through all the pages in one go.

INDEX

- 1 Project Purpose
- 2 Data Collection: *Web Scrapping*
- 3 Data Cleansing: *Regular expressions***
- 4 Pandas: *Final data frame + addition variables*
- 5 Key findings: *Using Tableau tables*
- 6 Next steps



Once we collected all the information through web-scraping it was time to start data cleansing. The product description contained a variety of information: brand, Mips technology etc. Using regular expressions, we extracted all this information into separate columns. Reminder, all steps were repreated for each website scrapped

1

Product description web-scraping output list

A screenshot of a web browser displaying three cycling helmets. The first is a 'Fox Racing Speedframe Pro MTB Helmet (MIPS)' at £139.00. The second is a 'POC Axion SPIN Helmet 2020' at £56.00 - £140.00. The third is a 'POC Tectal MTB Helmet Ed.) 2021' at £185.00. Below the browser is a Jupyter notebook cell labeled 'In [5]: names' which contains a list of helmet names.

```
In [5]: names
Out[5]: ['SixSixOne Reset Helmet',
          'IXS Trail EVO Helmet Exclusive 2020',
          'SixSixOne Reset MIPS Helmet',
          'Leatt MTB 2.0 Helmet 2021',
          'Fox Racing Dropframe Pro MTB Cycling Helmet',
          'Brand-X DH1 Full Face MTB Helmet AW20',
          'Fox Racing Speedframe Pro MTB Helmet (MIPS)',
          'POC Axion SPIN Helmet 2020',
          'POC Tectal MTB Helmet (Fabio Ed.) 2021',
          'Giro Fixture MTB Helmet (MIPS) 2019',
          'IXS Trigger FF Helmet 2019',
```

3

Cleaning prices column using regular expressions

A screenshot of a Jupyter notebook showing two code cells. The first cell, 'In [23]: price', displays a list of raw price strings including discounts like '£45.00\x00-\xa0\x00£89.99'. The second cell, 'In [24]: clean_prices', contains a script that uses regular expressions to find and replace these discount patterns with the base price.

```
In [23]: price
Out[23]: ['$45.00\x00-\xa0\x00£89.99',
           '$53.99',
           '$60.00\x00-\xa0\x00£119.99',
           '$47.99',
           '$180.00',
           '$59.99',
           '$139.00',
           '$56.00\x00-\xa0\x00£140.00',
           '$185.00',
           '$69.99',
           '$189.99\x00-\xa0\x00£199.99',
           '$160.00\x00-\xa0\x00£180.00',
           '$69.00']

In [24]: clean_prices = []
for i in price:
    try: clean_prices.append(re.findall(r"\d{1,3}.\d{2}",i)[0])
    except: clean_prices.append('0')
Out[24]: ['45.00',
           '53.99',
           '60.00',
           '47.99',
           '180.00',
           '59.99',
           '139.00',
           '56.00',
           '185.00',
           '69.99',
           '189.99',
           '160.00',
           '69.00']
```

2

Creating a brand column using regular expressions

A screenshot of a Jupyter notebook showing three code cells. The first cell, 'In [7]: brands_pattern', defines a regular expression pattern for brands like 'Fox Racing|POC|Troy Lee Designs|Bell|100'. The second cell, 'In [8]: brands', loops through names to find matches using this pattern. The third cell, 'In [22]: brands_clean', shows the resulting list of cleaned brand names.

- Creating a list of brands

```
In [7]: brands_pattern = re.compile(r"Fox Racing|POC|Troy Lee Designs|Bell|100")
In [8]: brands = []
for i in names:
    brands.append(re.findall(brands_pattern,i))
In [22]: brands_clean
Out[22]: ['SixSixOne',
          'IXS',
          'SixSixOne',
          'Leatt',
          'Fox Racing',
          'Brand-X',
          'Fox Racing',
          'POC',
          'POC',
          'Giro',
          'IXS',
          'POC',
```

4

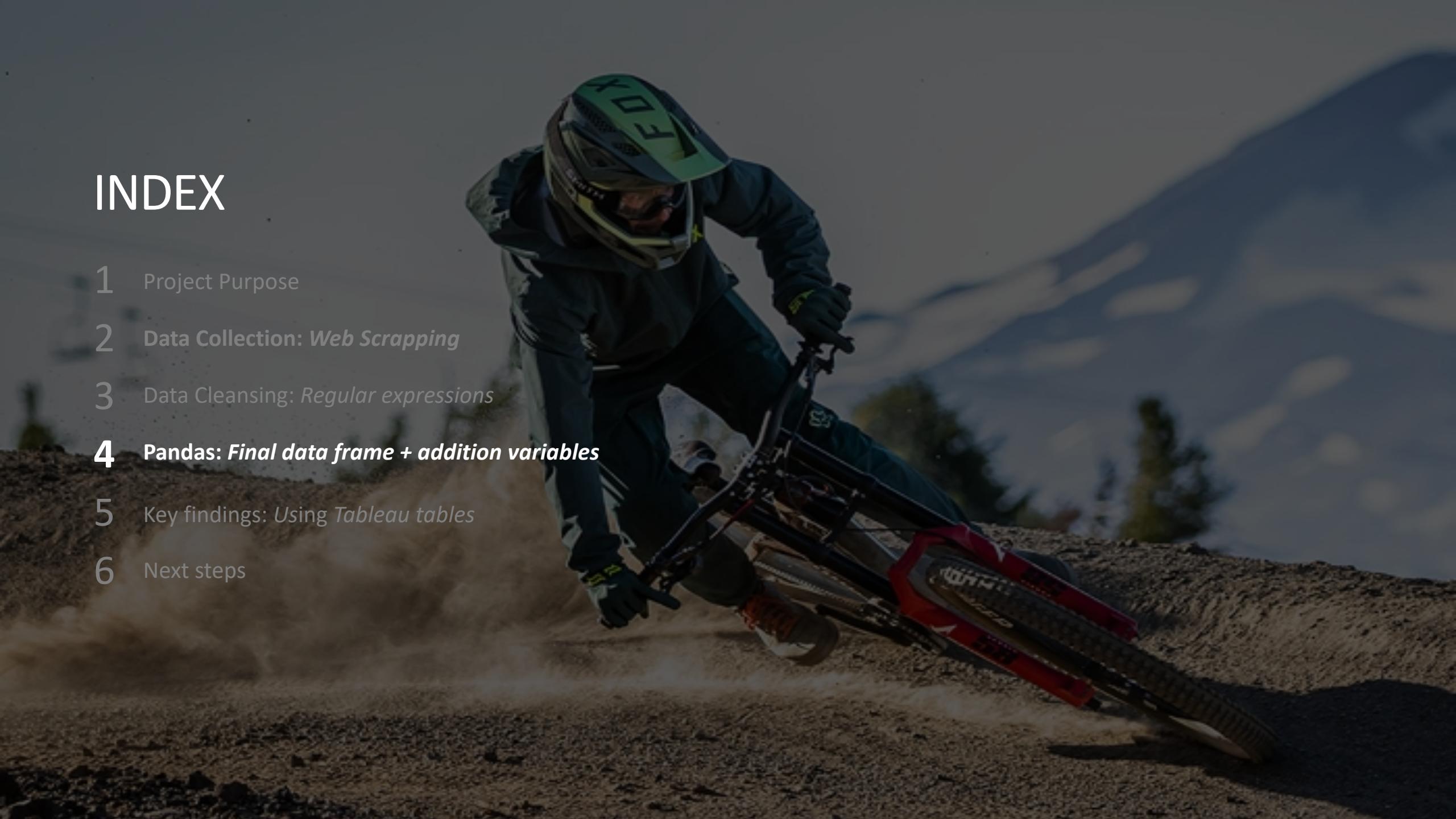
Creating a MIPS column using regular expressions

A screenshot of a Jupyter notebook showing two code cells. The first cell, 'In [28]: isMIPS', defines a regular expression pattern for 'MIPS'. The second cell, 'In [28]: isMIPS', loops through names to find matches using this pattern, resulting in a list where 1 indicates the helmet has MIPS and 0 indicates it does not.

- MIPS list

```
In [28]: isMIPS = []
pattern = re.compile(".*MIPS.*")
for name in product_descriptions:
    if pattern.match(name):
        isMIPS.append(1)
    else:
        isMIPS.append(0)
isMIPS
Out[28]: [0,
           0,
           1,
           0,
```

INDEX

- 
- 1 Project Purpose
 - 2 Data Collection: *Web Scrapping*
 - 3 Data Cleansing: *Regular expressions*
 - 4 Pandas: *Final data frame + addition variables***
 - 5 Key findings: *Using Tableau tables*
 - 6 Next steps

Using the pandas library, a data frame was constructed once the data cleansing was complete. After making sure all numerical values were indeed numerical, we now have a table to analyze.

Main Data Frame

```
In [74]: wiggle_dictionary = {"original_name": names,
                           "brand":brands_clean,
                           "item":product_description,
                           "price": clean_prices,
                           "oPrice": clean_oldprices,
                           "isMIPS" : isMIPS,
                           "year": years_clean2
                           }

wiggle = pd.DataFrame(wiggle_dictionary)
wiggle['oPrice'] = np.where(wiggle['oPrice'] == 0, wiggle['price'], wiggle['oPrice'])

wiggle[["price","oPrice"]] = wiggle[["price","oPrice"]].apply(pd.to_numeric)

In [75]: #Creating a scoring system:
wiggle["score_all_categories"] = (abs(len(wiggle) - np.array(wiggle.index))) / abs(len(wiggle))
```

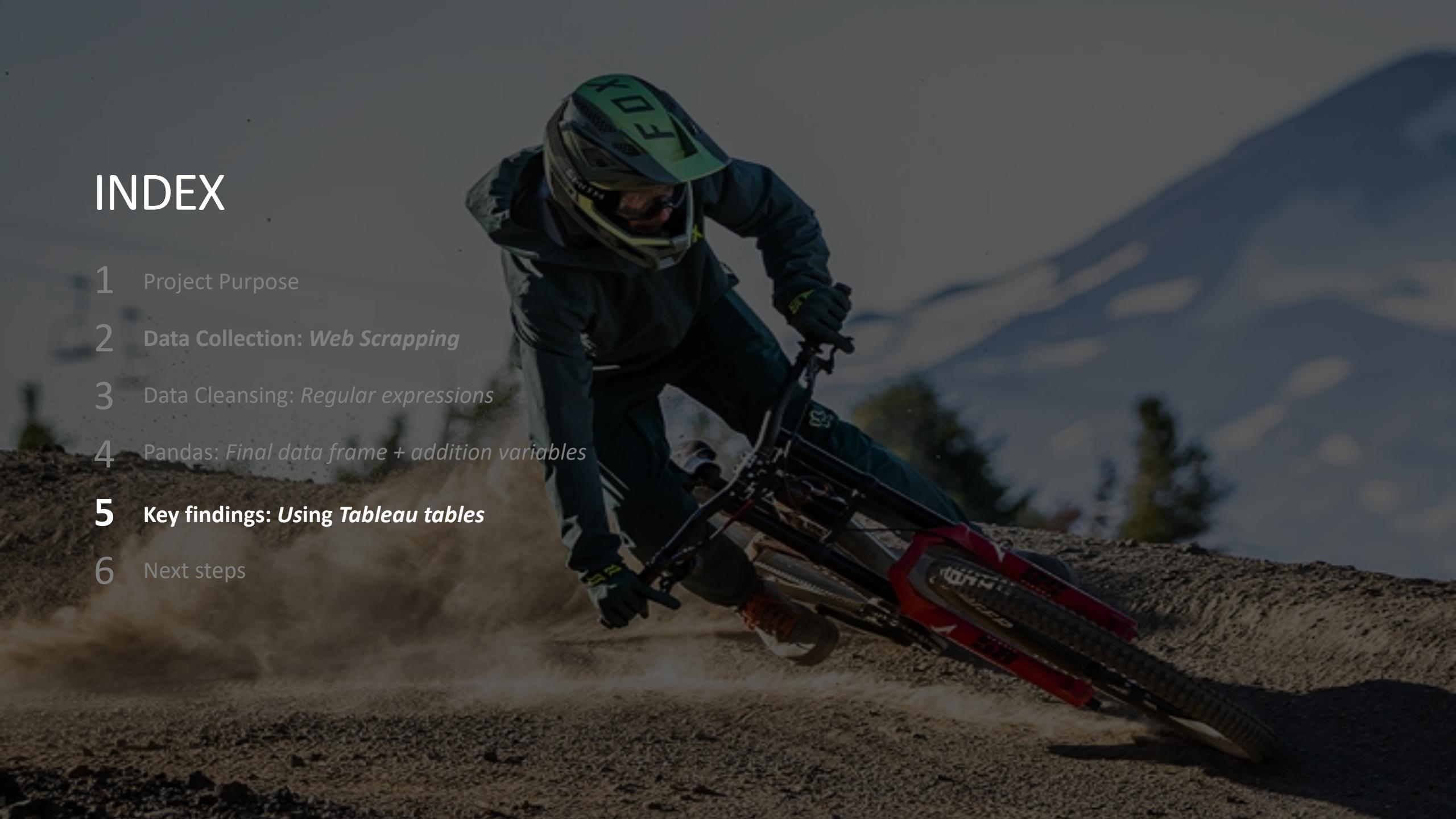
| | original_name | brand | item | price | oPrice | isMIPS | year | score_all_categories |
|-----|--|------------|---------------------------------|--------|--------|--------|-----------|----------------------|
| 0 | SixSixOne Reset Helmet | SixSixOne | Reset Helmet | 45.00 | 89.99 | 0 | Undefined | 1.000000 |
| 1 | IXS Trail EVO Helmet Exclusive 2020 | IXS | Trail EVO Helmet Exclusive 2020 | 53.99 | 89.99 | 0 | 2020 | 0.995633 |
| 2 | SixSixOne Reset MIPS Helmet | SixSixOne | Reset MIPS Helmet | 60.00 | 119.99 | 1 | Undefined | 0.991266 |
| 3 | Leatt MTB 2.0 Helmet 2021 | Leatt | MTB 2.0 Helmet 2021 | 47.99 | 79.99 | 0 | 2021 | 0.986900 |
| 4 | Fox Racing Dropframe Pro MTB Helmet AW20 | Fox Racing | Dropframe Pro MTB Helmet AW20 | 180.00 | 180.00 | 0 | 2020 | 0.982533 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 224 | Endura MT500JR Youth Helmet | Endura | MT500JR Youth Helmet | 74.99 | 74.99 | 0 | Undefined | 0.021834 |
| 225 | Leatt DBX 6.0 Carbon Helmet | Leatt | DBX 6.0 Carbon Helmet | 399.00 | 399.00 | 0 | Undefined | 0.017467 |
| 226 | IXS Kronos Evo MTB Helmet 2017 | IXS | Kronos Evo MTB Helmet 2017 | 74.99 | 74.99 | 0 | 2020 | 0.013100 |
| 227 | 7 iDP Project 23 Helmet Visor 2020 | 7 iDP | Project 23 Helmet Visor 2020 | 11.00 | 34.99 | 0 | 2020 | 0.008734 |
| 228 | Leatt DBX 2.0 Helmet 2020 | Leatt | DBX 2.0 Helmet 2020 | 79.99 | 79.99 | 0 | 2020 | 0.004367 |

Visibility Score

The visibility score was constructed to measure where each product is positioned on the webpage. The main reason behind this is that the product which is shown first has a higher probability of being seen (and thus bought) in comparison to the last products on the page.

Thus, since the web-scrap took products in the order they appear on the website, we used the index to create a point system from the score between 0 & 1. In this case, 1 is the first product that appears and 0 is the last product that appears.

INDEX

- 
- 1 Project Purpose
 - 2 Data Collection: *Web Scrapping*
 - 3 Data Cleansing: *Regular expressions*
 - 4 Pandas: *Final data frame + addition variables*
 - 5 Key findings: *Using Tableau tables*
 - 6 Next steps

5.1 Key Findings on generic MTB helmets page in CRC

(Contains all MTB helmet products regardless of sub-categories)



Chain Reaction Cycles FOR EVERY RIDE

Search by keyword, product or brand...

Sign In | My Account | Help | Wishlist | GBP £GBP

SHOP BY CATEGORY | BRANDS | MTB | ROAD | CITY | BMX | INDOOR TRAINING | CLEARANCE

BLACK FRIDAY

Home > Protection > Helmets

YOUR CHOICES

MTB Clear All

NARROW YOUR RESULTS

Categories

Protection

Helmets

- Dirt Jump Helmets (4)
- Full Face Helmets (56)
- Helmets Spares (44)
- Kids Helmets (10)
- MTB Helmets (115)

Price (£)

from _____ to _____ Go

Status

In Stock Only (181)

View 24 48 72

Sort By New

Black Friday

SixSixOne Reset Helmet

£45.00 - £89.99 RRP £89.99 SAVE UP TO 50%

IXS Trail EVO Helmet Exclusive 2020

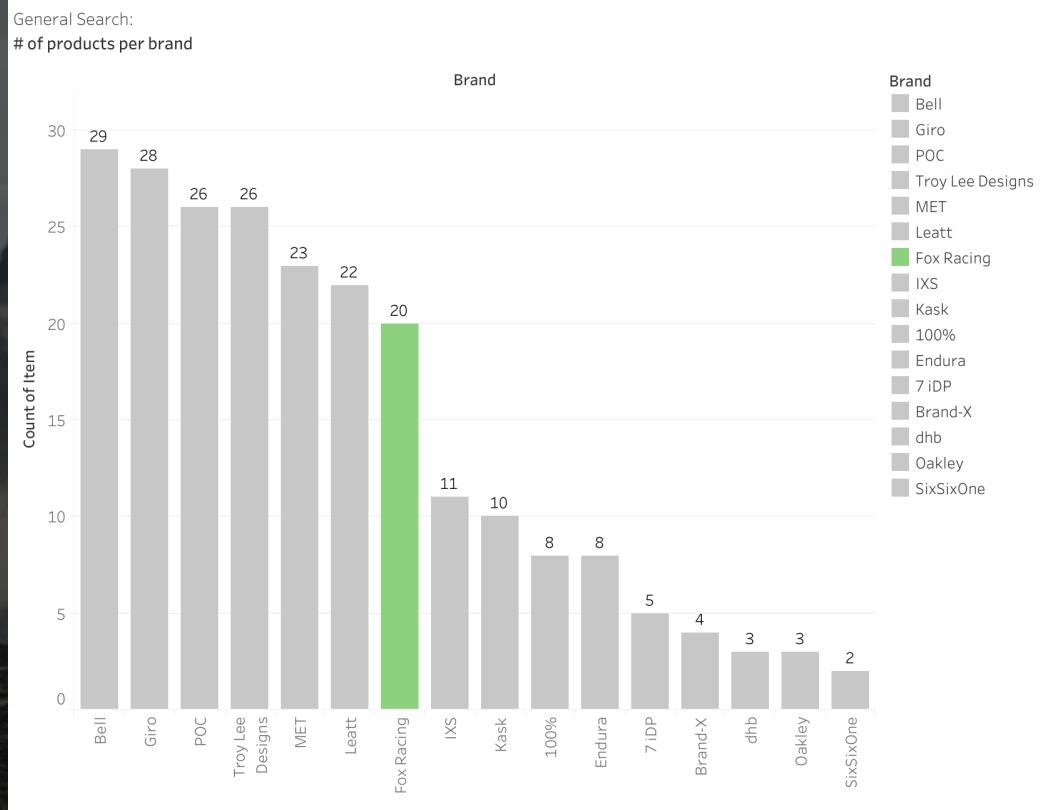
£53.99 RRP £89.99 SAVE 40%

SixSixOne Reset MIPS Helmet

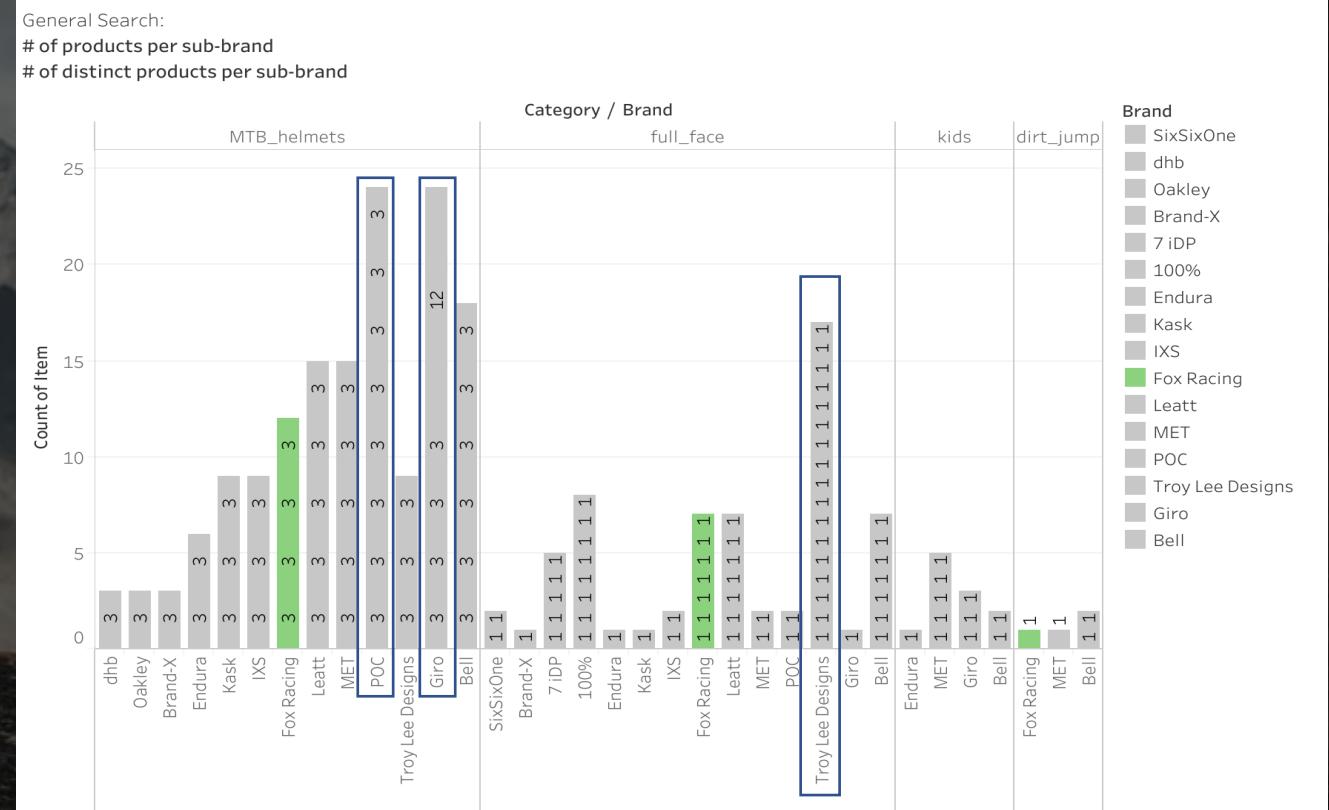
£60.00 - £119.99 RRP £119.99 SAVE UP TO 50%

This page includes all products related to MTB helmets. This includes all the subcategories inside MTB helmets: dirt-jump; full-face, kids & MTB helmets.

Fox Racing has 20 unique products represented in the general category page. In this page, we can see a large representation in subcategories MTB helmet & full face helmets. However, POC & Giro lead product representation in MTB helmets and Troy-Lee Designs leads in full-face helmets



The following data is based on the generic MTB Helmet page

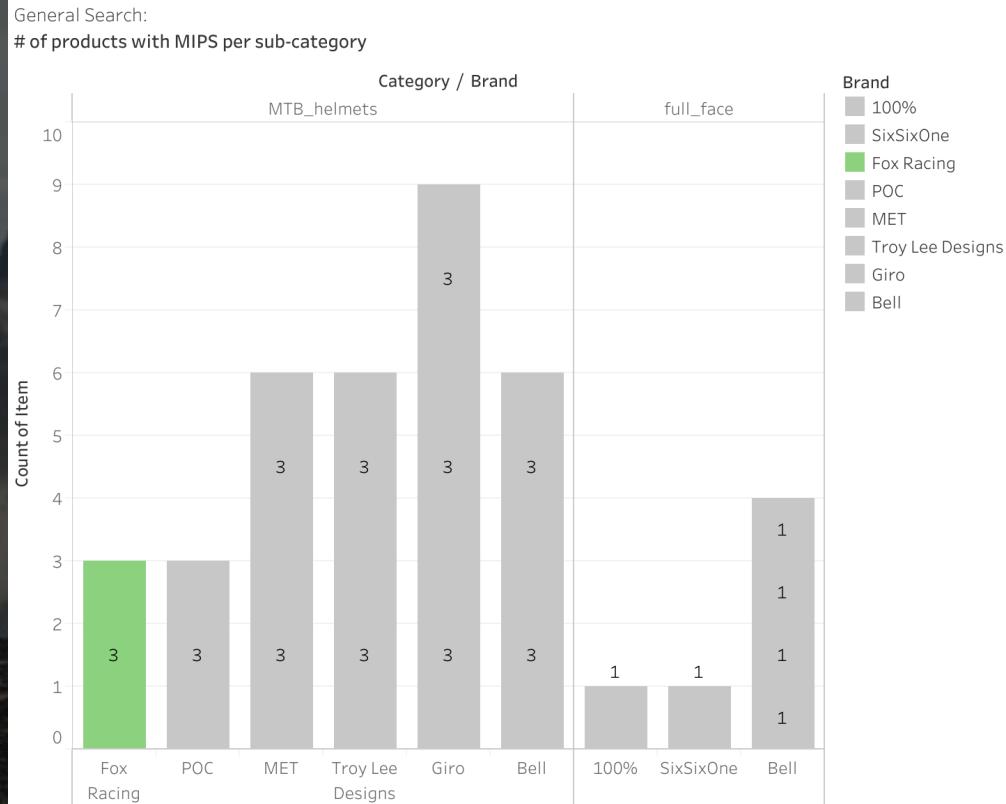


The following data is based on the main MTB Helmet pages

- **How to read the #'s inside the bars – example:** Fox Racing has 4 distinct products in the MTB-helmet category and each of these products are represented 3 times inside the page.
- **Insight –** MTB Helmet sub-category has each distinct product represented on average 3x on the page. While full-face has one of each product. There is also an opportunity for brand representation in the kids category.



When looking into products descriptions that contain MIPS technology, we can see that Fox Racing has only one product. In addition, it does not have any representation in the full face helmets.



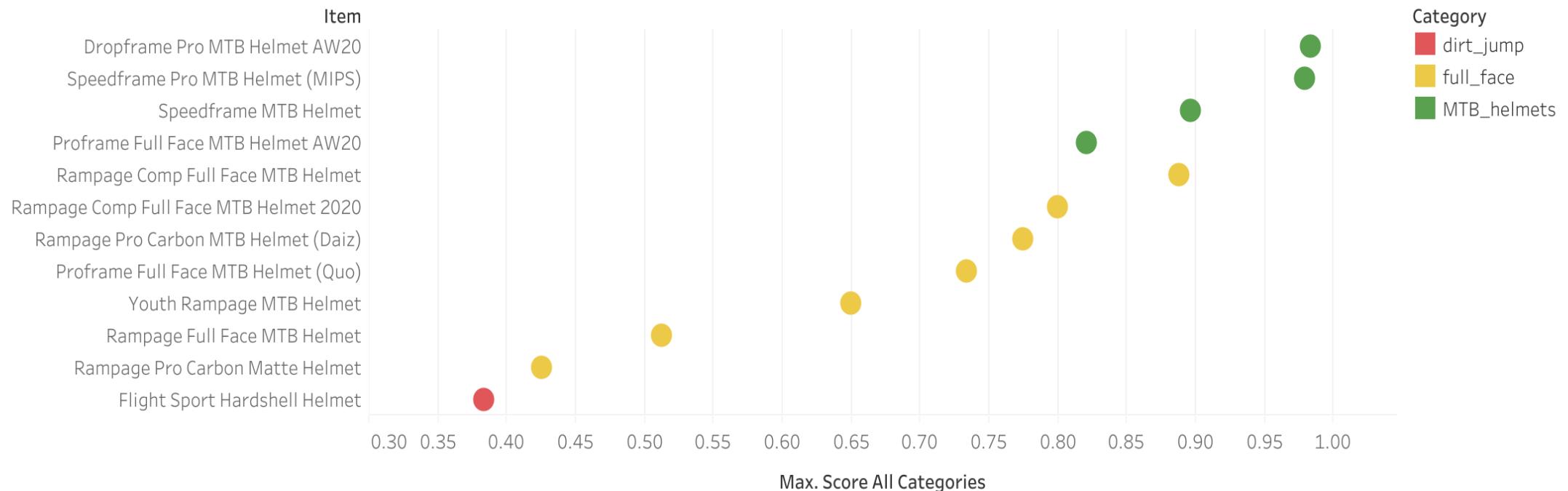
The following data is based on the main MTB Helmet pages



Fox Racing products show a high visibility score on the general page with MTB helmets, while full-face helmets and dirt-jump category products are not well positioned in the general MTB helmets page.

General Search:

Fox Racing visibility scores

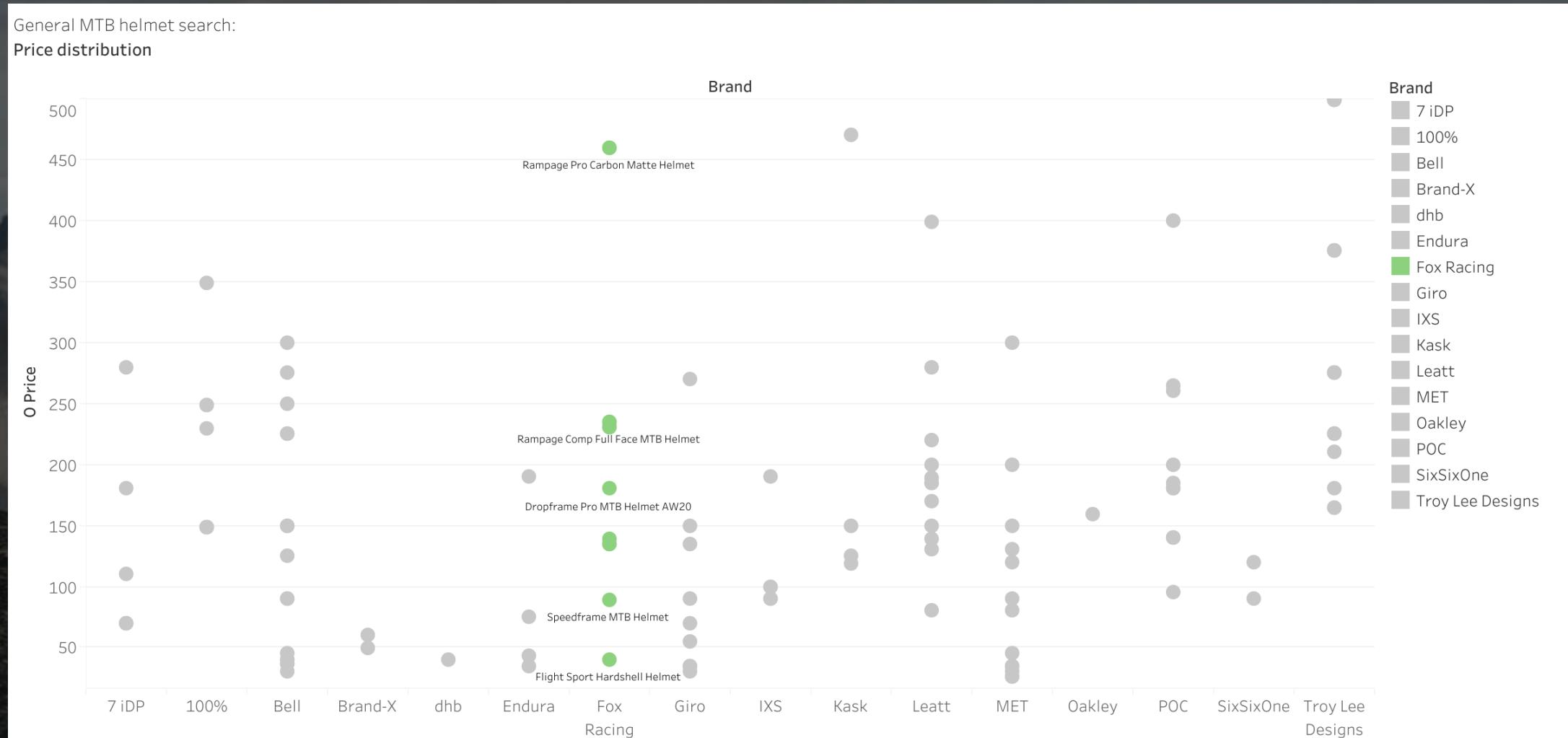


The following data is based on the main MTB Helmet pages

- Explaining the graph:** A visibility score was created to calculate the product positioning on the web page. In this case, if the visibility score is exactly one, this signifies that it is the first product to appear on the webpage. Although some products are repeated on the webpage, we took the maximum visibility score for each distinct product



Fox Racing shows a spread-out price distribution between all its products. We can also see it has few distinct SKU's in comparison to other brands.



The following data is based on the main MTB Helmet pages

5.2 Key Findings on sub-category: MTB helmets pages in CRC

Search by keyword, product or brand... SEARCH

Sign In | My Account | Help | Wishlist | £GBP

SHOP BY CATEGORY | BRANDS | MTB | ROAD | CITY | BMX | INDOOR TRAINING | CLEARANCE

BLACK FRIDAY NEW DEALS ADDED EVERY THURSDAY WEEK 3

Home > Protection > Helmets

YOUR CHOICES

MTB X
MTB Helmets X

NARROW YOUR RESULTS

Categories

- Protection
- Helmets
- MTB Helmets

Price (£)

from _____ to _____ Go

Status

In Stock Only (87)

Gender

Female (103)
 Male (115)

MTB Helmets

View 24 48 72 grid

Showing 1 - 48 of 115 Sort By New

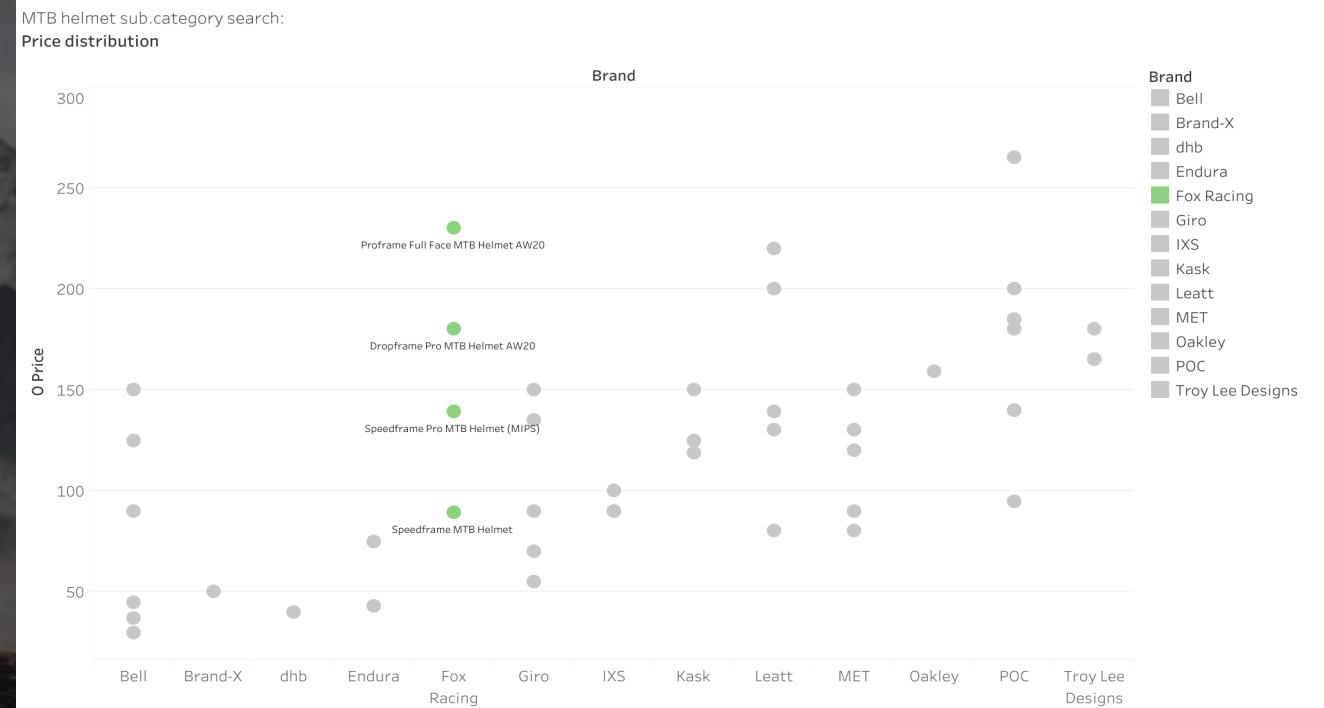
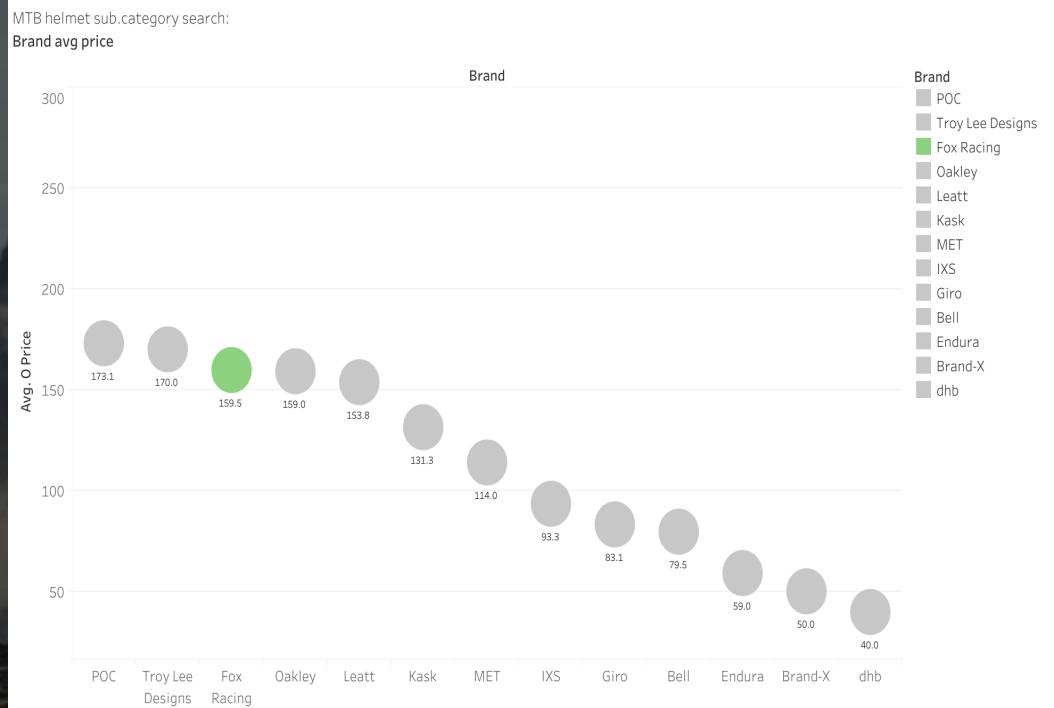
BLACK FRIDAY **BLACK FRIDAY** **BLACK FRIDAY**

IXS Trail EVO Helmet Exclusive 2020
£53.99 ~~£89.99~~ SAVE 40%

Leatt MTB 2.0 Helmet 2021
£47.99 ~~£79.99~~ SAVE 40%

Fox Racing Dropframe Pro MTB Helmet AW20
£180.00

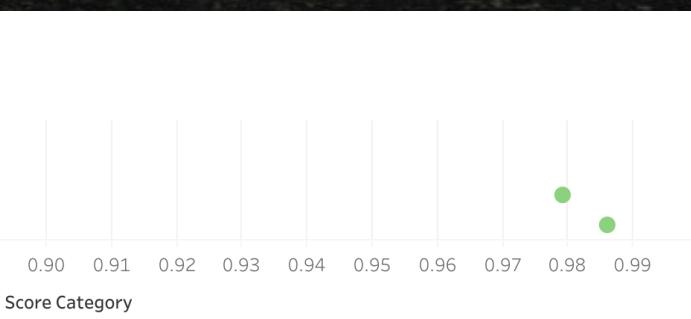
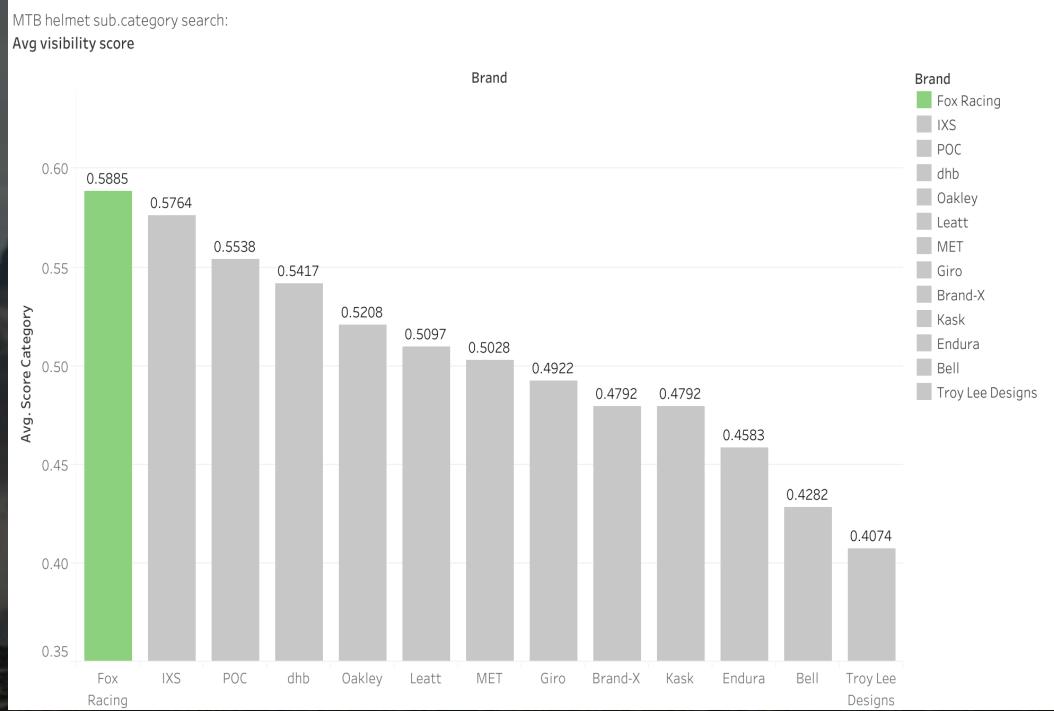
Fox Racing in the MTB-helmet subcategory is on average a high priced brand, with only POC & Troy Lee Designs with a higher average price. In terms of price distribution, Fox Racing is well distributed.



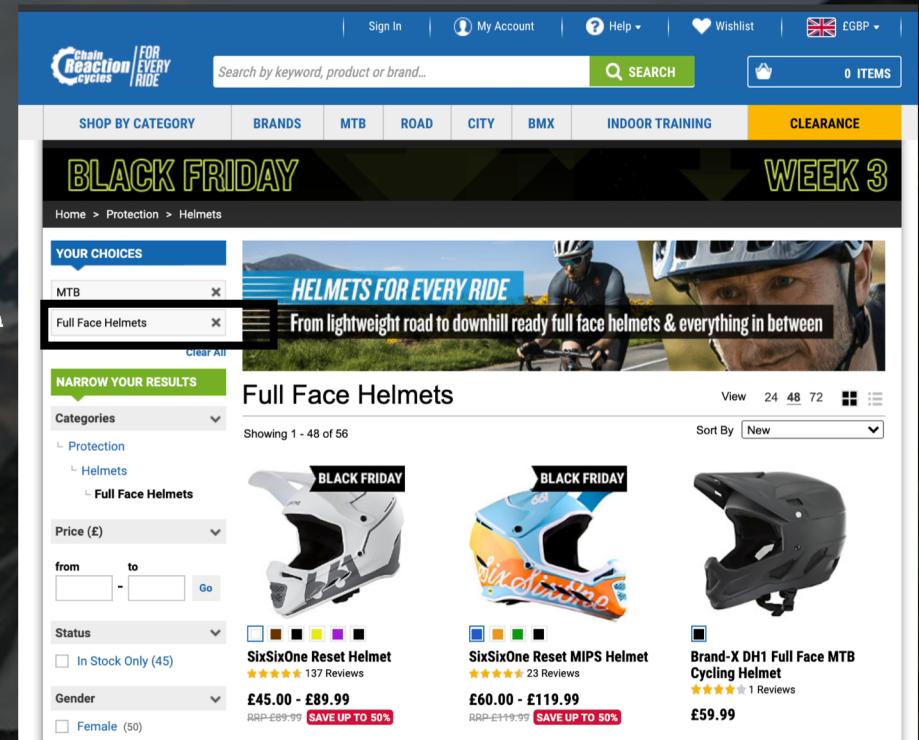
The following data is based on the sub category MTB Helmet page

The following data is based on the sub category MTB Helmet page

Fox Racing is the leading brand in terms of visibility in the CRC website inside the MTB helmet subcategory page. This is great news as it is likely to result in more revenue.

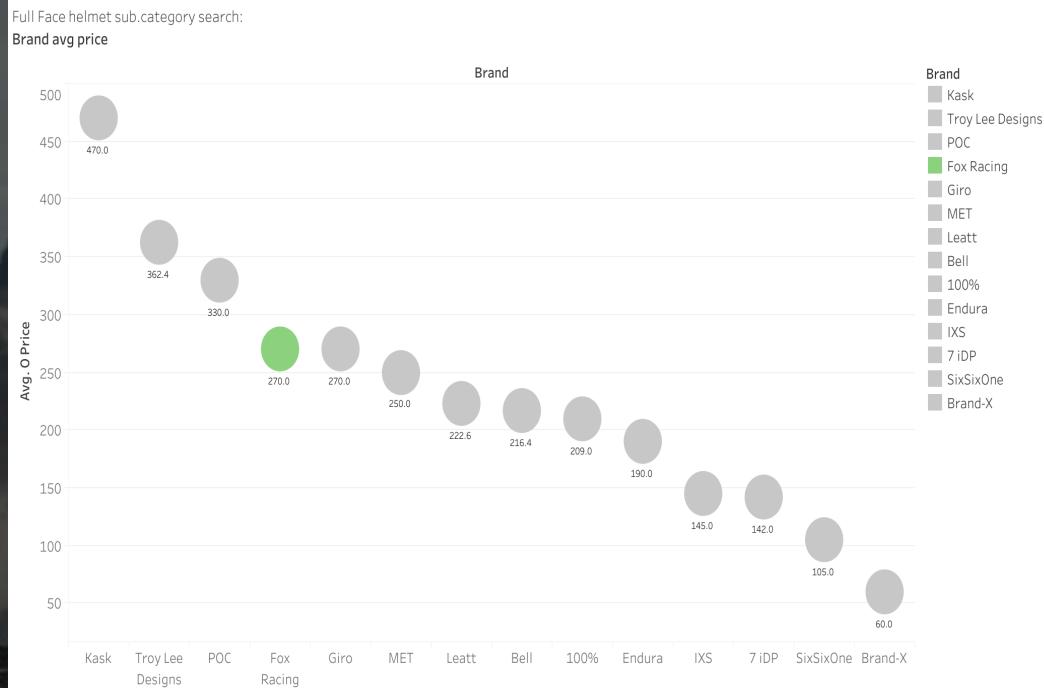


5.3 Key Findings on sub-category: Full-Face helmets pages in CRC



The screenshot shows the Chain Reaction Cycles website interface. At the top, there's a navigation bar with links for Sign In, My Account, Help, Wishlist, and a currency selector (GBP). Below the header is a search bar and a 'SEARCH' button. The main menu includes categories like SHOP BY CATEGORY, BRANDS, MTB, ROAD, CITY, BMX, INDOOR TRAINING, and CLEARANCE. A prominent 'BLACK FRIDAY' banner is displayed across the top. The current page is 'Home > Protection > Helmets'. On the left, a sidebar titled 'YOUR CHOICES' shows 'MTB' and 'Full Face Helmets' selected. Below it, 'NARROW YOUR RESULTS' allows filtering by Categories (Protection > Helmets > Full Face Helmets), Price (£), Status (In Stock Only), and Gender (Female). The main content area displays three full-face helmets: a white SixSixOne Reset Helmet, a blue and orange SixSixOne Reset MIPS Helmet, and a black Brand-X DH1 Full Face MTB Cycling Helmet. Each item has a 'BLACK FRIDAY' badge, its name, a star rating, the original price (e.g., £89.99), the discounted price (e.g., £45.00 - £89.99), and a 'SAVE UP TO 50%' offer.

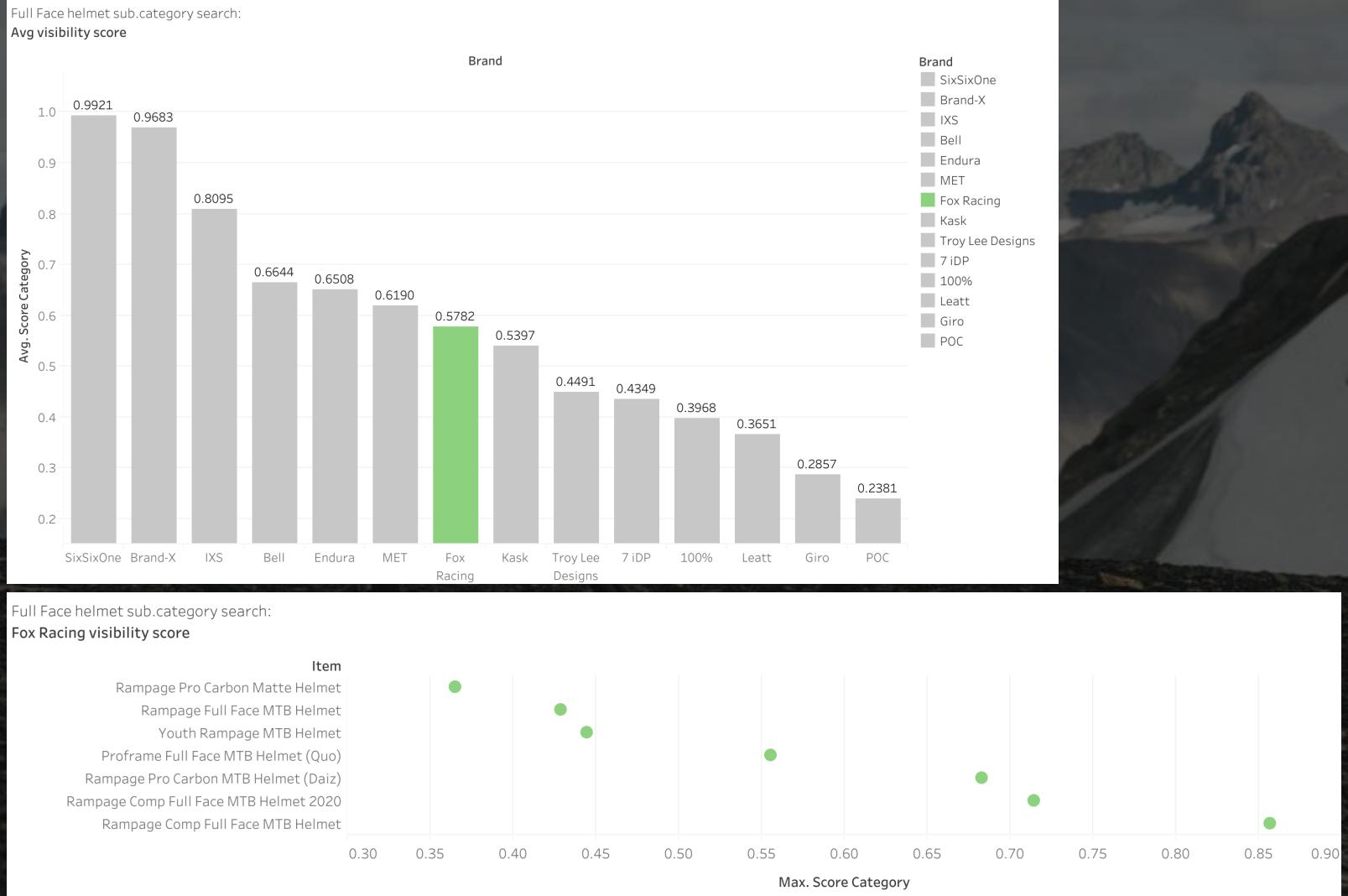
Fox Racing in the Full-Face-helmet subcategory is on average a high priced brand with the same average price as Giro. We can see that in this category, Fox Racing has on avg fewer distinct SKU's.



The following data is based on the sub category MTB Helmet page

The following data is based on the sub category MTB Helmet page

Fox Racing has a poor visibility score in the Full-faced helmets category. Especially if comparing these results to the subcategory of MTB helmets we saw previously.

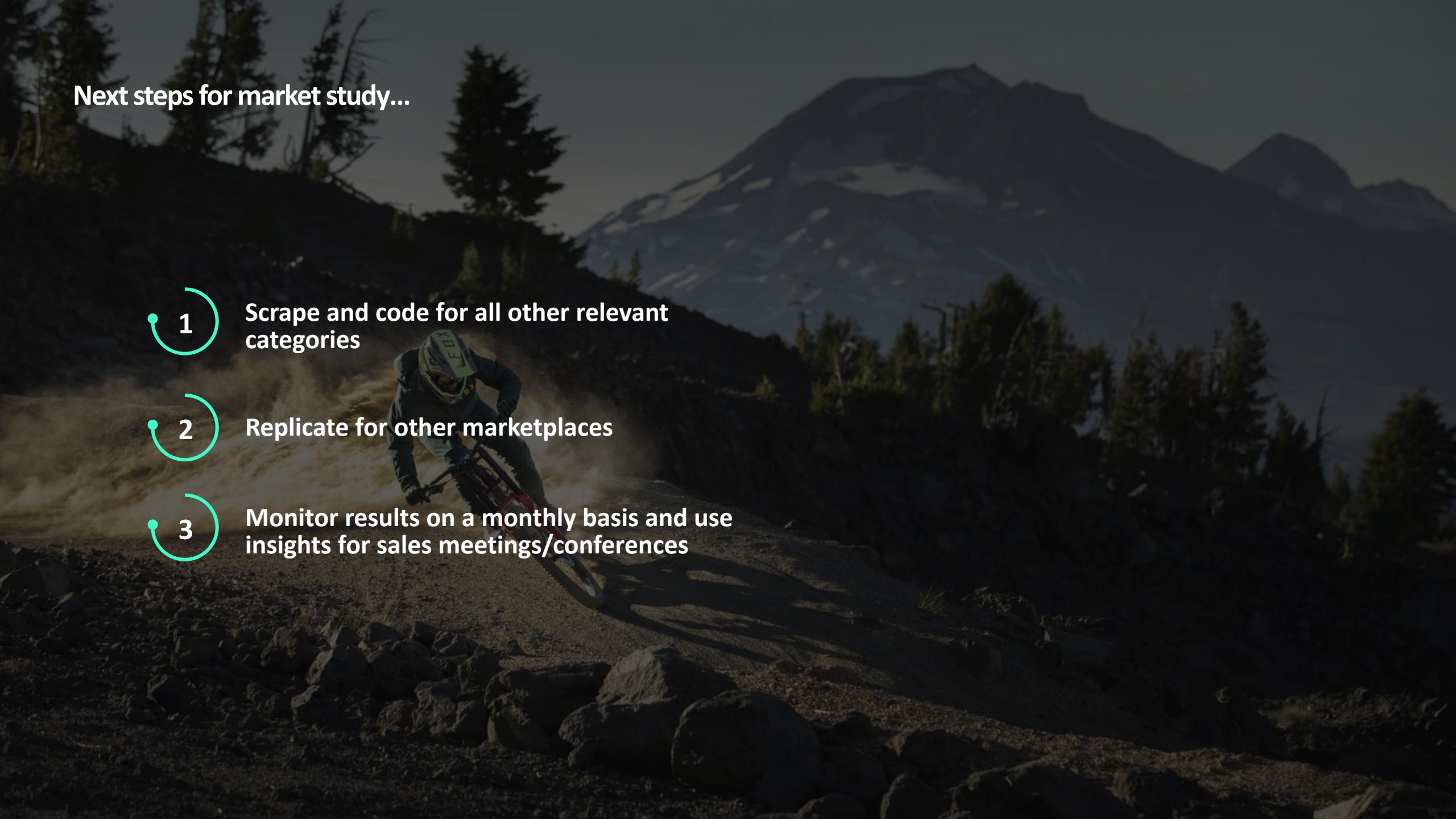


INDEX

- 1 Project Purpose
- 2 Data Collection: *Web Scrapping*
- 3 Data Cleansing: *Regular expressions*
- 4 Pandas: *Final data frame + addition variables*
- 5 Key findings: *Using Tableau tables*
- 6 Next steps



Next steps for market study...

- 
- A photograph of a mountain biker in mid-air, performing a jump on a rocky trail. The biker is wearing a green helmet and a dark jacket with 'FD' on the back. The background features a majestic mountain range with snow-capped peaks under a clear sky.
- 1 Scrape and code for all other relevant categories
 - 2 Replicate for other marketplaces
 - 3 Monitor results on a monthly basis and use insights for sales meetings/conferences



Technical Report

Mountain Bike Helmets Marketplace Study using Chain Reaction Cycles

11/23/2020

André Santa Clara