



HIGHER SCHOOL OF ECONOMICS

POWER BI PRESENTATION



END OF COURSE PROJECT
PRESENTATION BY NGOYE MICHEL



HIGHER SCHOOL OF ECONOMICS



Presentation outline

Key points:

- Presentation of the dataset
- Executive Summary
- Understanding our customers
- Understanding our products
- How are we doing from a sales perspective?

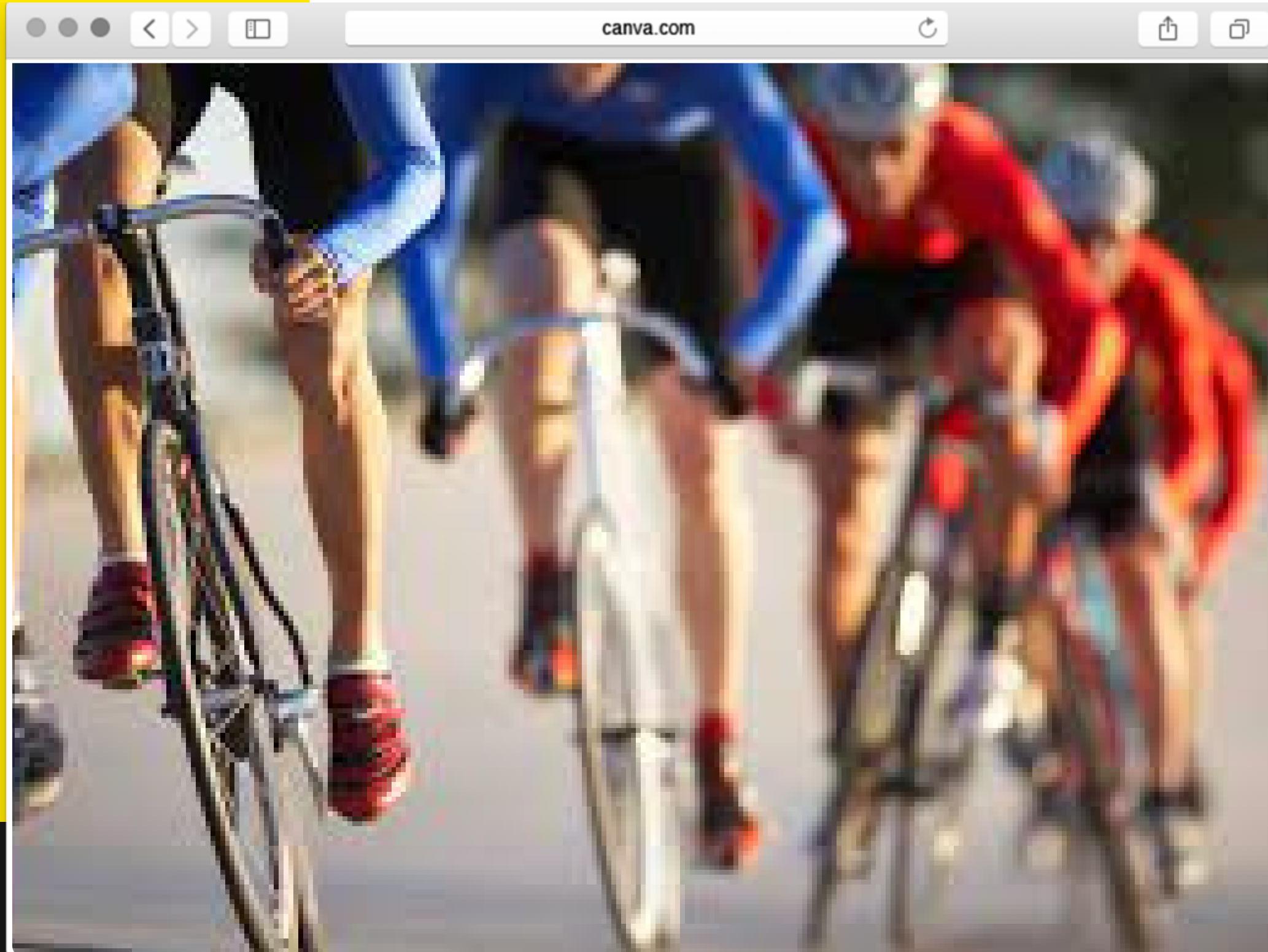


AdventureWorks cycles

**Large multinational
manufacturing company**

What does the company produces?

- metal and composite bikes
- bike accessories
- components
- clothing





Companies presence and channels of distribution

The company operates in North America, Europe, Asia and Australia

The main channels of distribution are resellers and the internet

Presentation of the dataset

Customer table



contains information about
customers

Date table



Containing information about fiscal
year, dates, quarters, months and
datekeys

Product table



containing information about the
products, information such as
names, standard cost, list price,
categories and subcategories

Reseller table



the company is also selling to
resellers and this contains
information about them, their
location and the type of business
they operate

**Sales and Sales
Order Tables**



which is the fact table from where
all information will be derived.

Queries [9]

= Table.TransformColumnTypes(Sales_Table,{{"SalesOrderLineKey", Int64.Type}, {"ResellerKey", Int64.Type}, {"CustomerKey", Int64.Type}, {"ProductKey", Int64.Type}, {"OrderDateKey", Int64.Type}, {"DueDateKey", Int64.Type}, {"ShipDateKey", Int64.Type}})

	SalesOrderLineKey	ResellerKey	CustomerKey	ProductKey	OrderDateKey	DueDateKey	ShipDateKey
1	43659001	676	-1	349	20170702	20170712	
2	43659002	676	-1	350	20170702	20170712	
3	43659003	676	-1	351	20170702	20170712	
4	43659004	676	-1	344	20170702	20170712	
5	43659005	676	-1	345			
6	43659006	676	-1	346			
7	43659007	676	-1	347			
8	43659008	676	-1	229			
9	43659009	676	-1	235			
10	43659010	676	-1	218			
11	43659011	676	-1	223			
12	43659012	676	-1	220			
13	43660001	117	-1	326			
14	43660002	117	-1	319			
15	43661001	442	-1	300			
16	43661002	442	-1	296			
17	43661003	442	-1	304			
18	43661004	442	-1	223			
19	43661005	442	-1	232			
20	43661006	442	-1	293			
21	43661007	442	-1	348			
22							

15 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows

Query Settings

PROPERTIES

Name: Sales
All Properties

APPLIED STEPS

- Source
- Navigation
- Changed Type

Data cleaning

The data was very clean, probably due to the fact that it's from fictitious company created by Microsoft. Few transformations had to be made still:

- changing the data type for Power BI to function properly
- a few top rows also had to be removed

AdventureWorks_Michel - Power Query Editor

File Home Transform Add Column View Tools Help

Close & Apply New Source Data source settings Manage Parameters Refresh Preview Advanced Editor Choose Columns Remove Columns Keep Rows Remove Rows Sort Data Type: Whole Number Data Merge Queries Use First Row as Headers Group By Split Column Replace Values Append Queries Combine Files Combine Files Azure Machine Learning Close New Query Data Sources Parameters Query Manage Columns Reduce Rows Sort Transform Combine AI Insights

Queries [9] X ✓ fx

= Table.RenameColumns(#"Expanded Sheet1 (2)",{{"Sex", "Gender"}})

	State-Province	Country-Region	Postal Code	Gender	Age	IncomeLevel
1	Queensland	Australia	4700	F	38	Average
2	Victoria	Australia	3198	F	30	Very High
3	Tasmania	Australia	7001	M	45	Low
4	New South Wales	Australia	2112	M	57	Low
5	New South Wales	Australia				
6	Queensland	Australia				
7	New South Wales	Australia				
8	Victoria	Australia				
9	Victoria	Australia				
10	Queensland	Australia				
11	Queensland	Australia				
12	Queensland	Australia				
13	Washington	United States				
14	Oregon	United States				
15	Washington	United States				
16	California	United States				
17	California	United States				
18	Victoria	Australia				
19	Victoria	Australia				
20	British Columbia	Canada				
21	British Columbia	Canada				
22						

Feature engineering

As for the customer's table, a few features that were not available had to be created in Excel and then merged to the table using the merge queries tool. The gender, Age and Income level columns were merged.



AutoSave Off

B2 =CHOOSE(RANDBETWEEN(1,2),Sheet2!\$A\$1,Sheet2!\$A\$2)

Feature engineering

To engineer the Gender column we used the formula

CHOOSE(RANDBETWEEN(1,2), SOME
PREDEFINED VALUES)

As for the Age feature we used the formula

RANDBETWEEN(19,60)

HIGHER SCHOOL OF ECONOMICS

File Home Insert Page Layout Formulas Data Review View Help

Cut Copy Paste Format Painter

Font Alignment Number

Clipboard

AutoSave Off

B2 : fx =CHOOSE(RANDBETWEEN(1,4),Sheet2!\$A\$1,Sheet2!\$A\$2,Sheet2!\$A\$3,Sheet2!\$A\$4)

	A	B	C	D	E	F	G	H	I	J	K
1	CustomerKey	IncomeLevel									
2	-1	High									
3	11000	Low									
4	11001	Low									
5	11002	Very High									
6	11003	Very High									
7	11004	Average									
8	11005	Low									
9	11006	High									
10	11007	Average									
11	11008	High									
12	11009	Very High									
13	11010	High									
14	11011	Low									
15	11012	Average									
16	11013	High									
17	11014	Average									
18	11015	Average									
19	11016	Very High									
20	11017	High									
21	11018	Low									
22	11019	High									
23	11020	Low									
24	11021	Low									
25	11022	Average									
26	11023	Average									
27	11024	Very High									

Sheet1 Sheet2

Calibri 11 A A General \$ % ,

Cut Copy Paste Format Painter Merge & Center Wrap Text

Clipboard Font Alignment

B8 : fx

	A	B	C	D	E	F	G	H	I	J
1	Average									
2	High									
3	Low									
4	Very High									
5										
6										

Feature engineering

As for the income level column we used the formula
=CHOOSE(RANDBETWEEN(1,4),Sheet2!
\$A\$1,Sheet2!\$A\$2,Sheet2!\$A\$3,Sheet2!
\$A\$4)
to generate values

HIGHER SCHOOL OF ECONOMICS

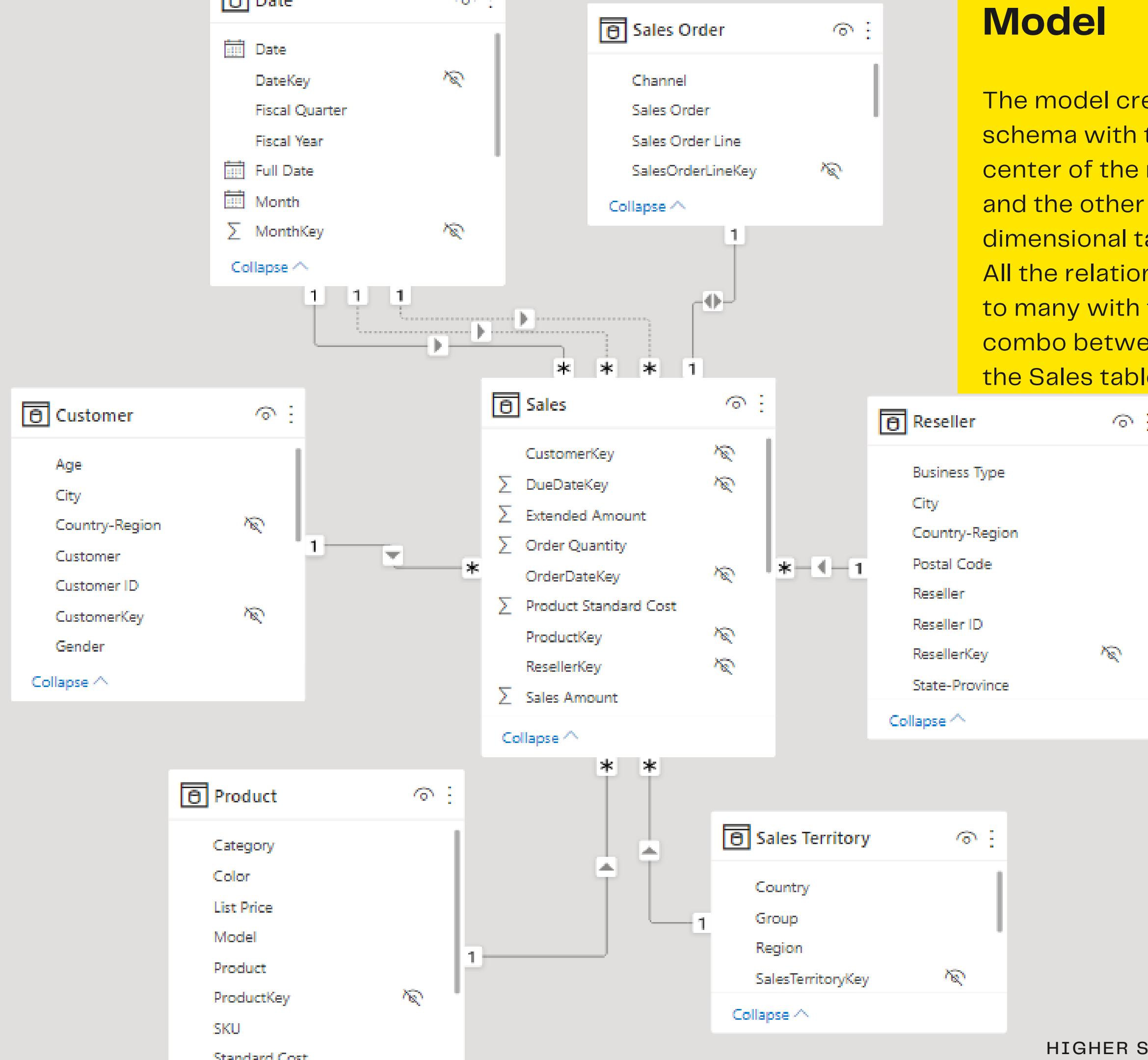
Type here to search

Model

The model created is a traditional star schema with the Sales table at the center of the model as a fact table,

and the other 6 tables acted as dimensional tables

All the relationships types were one to many with triple relationship combo between the data table and the Sales table





canva.com

Power BI Measures

New Column Table New Parameter Sort by Column

Avg_Qty_Sold = AVERAGE('Sales Data'[Quantity])

Product Type	Quantity	Avg_Qty_Sold
Tablet	1232	64.00
Laptop	1320	39.33
Desktop	3094	26.34
Laptop	1330	37.50
Tablet	2220	44.40
Laptop	2396	43.56
Mobile	2136	42.12
Laptop	5202	41.29
Mobile	1094	40.60

Data type: Decimal Number
Format: General
Size: No. of digits: Auto
Home Table: Sales Data
Data Category: Default
Default Unit:

Avg_Qty_Sold by City Name

Some measures were created



Se

Help

Table tools

Measure tools



Format General



\$ % , .00 Auto



Data category



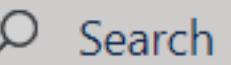
Uncategorize

Properties

Formatting

Total Profits = [Total revenue] - [Total Costs]

	ResellerKey	CustomerKey	ProductKey	OrderDateKey	DueDateKey
3001	510	-1	322	20170707	20170707



Search

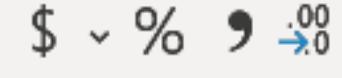
Help

Table tools

Measure tools



Format General



\$ % , .00 Auto



Data category Uncategorized



New measure



Quick measure

Properties

Calculations

Total Costs = SUMX(Sales, Sales[Order Quantity]* RELATED('Product'[Standard Cost]))

SKU	Product	Standard Cost	Color	List Price	Model	Subcategory	Category
FR-R92B-58	HL Road Frame - Black, 58	868.6342	Black	1431.5	HL Road Frame	Road Frames	Components

Feature engineering

As for the income level column we used the formula

```
=CHOOSE(RANDBETWEEN(1,4),Sheet2!$A$1,Sheet2!$A$2,Sheet2!$A$3,Sheet2!$A$4)
```

to generate values

HIGHER SCHOOL OF ECONOMICS

Help Table tools Measure tools

revenue

\$% Format General Data category Uncategorized

\$ % , .00 Auto

Formatting Properties

Total revenue = `SUMX(Sales, Sales[Order Quantity]*RELATED('Product'[List Price]))`

ResellerKey	CustomerKey	ProductKey	OrderDateKey	DueDateKey
563001	510	-1	322	20170707
566001	511	-1	330	20170709

AdventureWorks_Michel - Power BI Desktop

Search

Help Table tools Measure tools

Orders

\$% Format Decimal number Data category Uncategorized

\$ % , .00 0

Formatting Properties

Total Orders = `DISTINCTCOUNT('Sales Order'[SalesOrderLineKey])`

SalesOrderLineKey	Sales Order	Sales Order Line
53465001	SO53465	SO53465 - 1
53465002	SO53465	SO53465 - 2
53465003	SO53465	SO53465 - 3

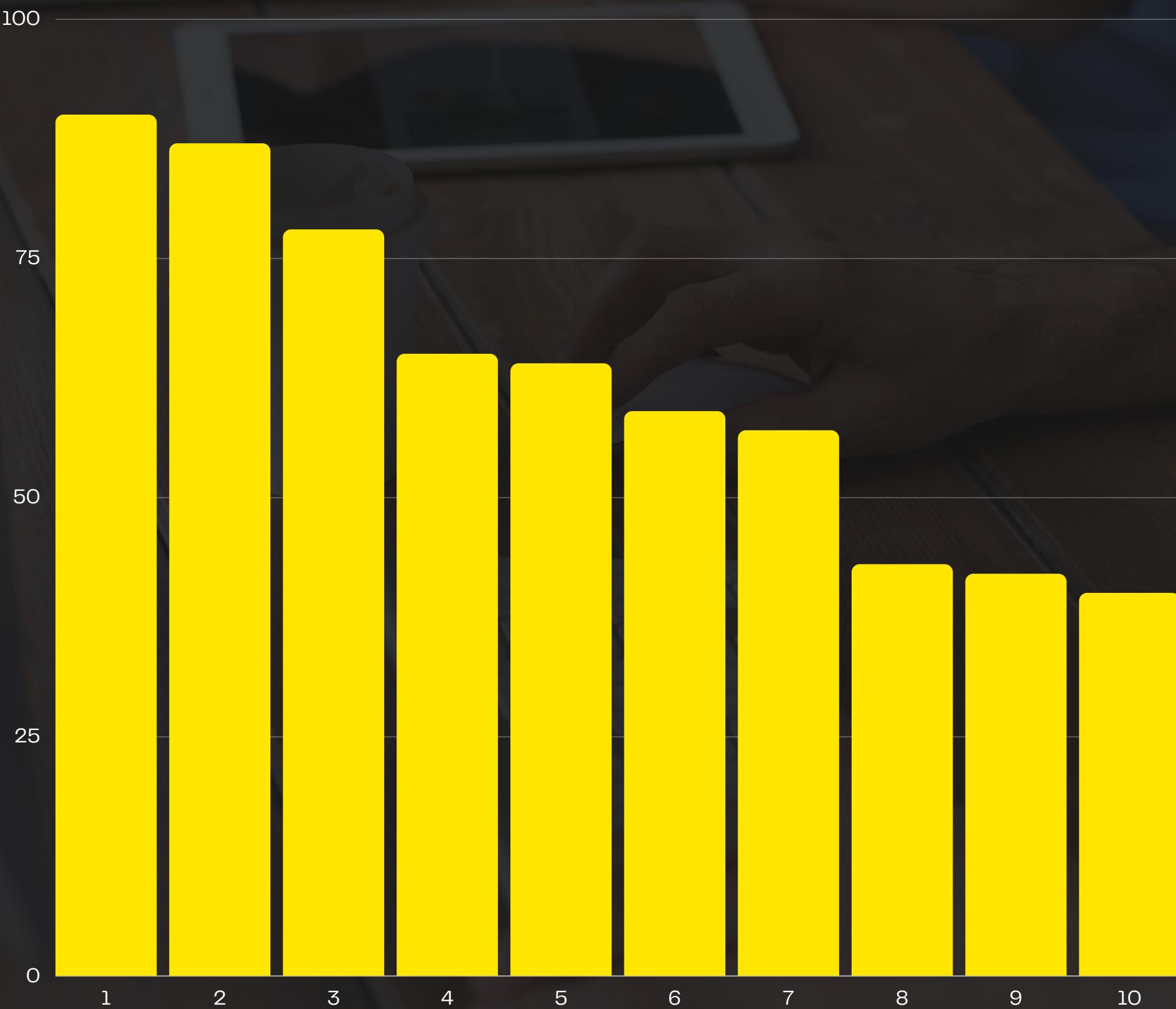
Feature engineering

As for the income level column we used the formula
`=CHOOSE(RANDBETWEEN(1,4),Sheet2!A1,Sheet2!A2,Sheet2!A3,Sheet2!A4)`
to generate values

5

Dashboards created

- Executive summary
- Customer Analysis
- Reseller's analysis
- Product Analysis
- Sales Analysis



Executive Summary -- Sales Report



Calendar

2019

Region

total Sales

2.90M

Total Costs

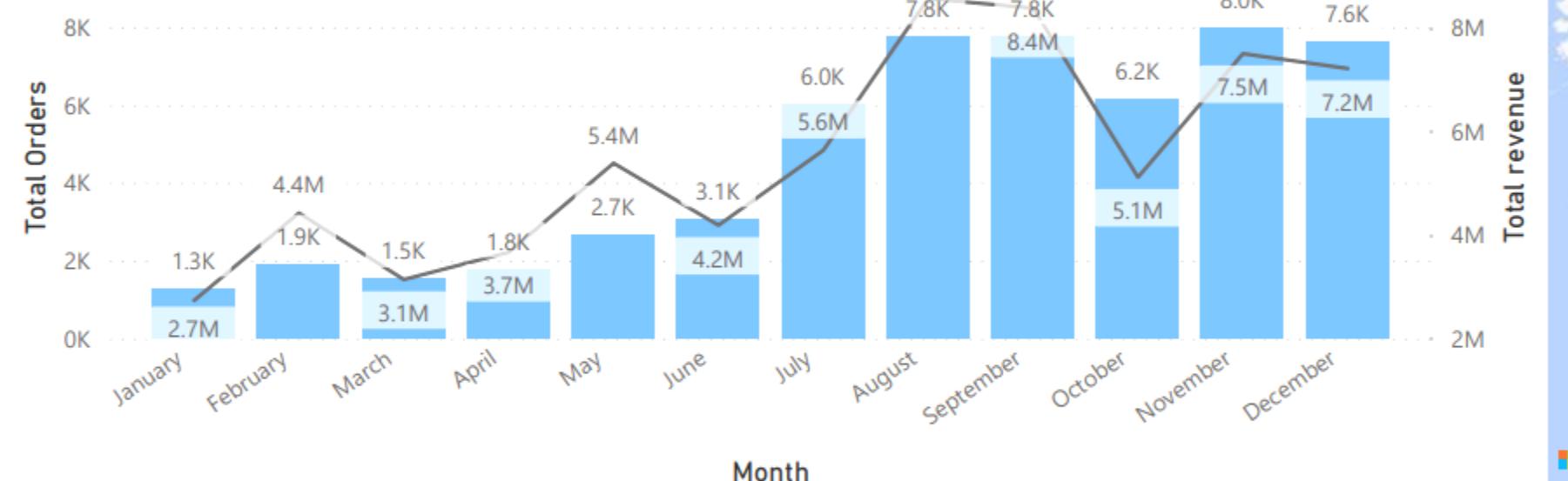
3.62M | 27.36M

Profits

36M | 56K

How much are we selling and how does it affect revenues ?

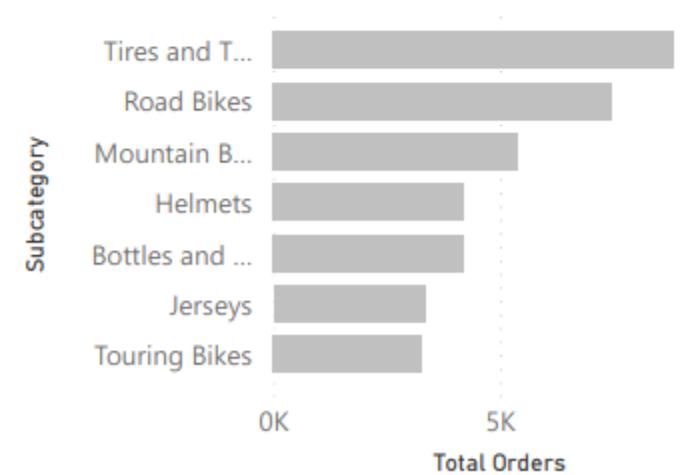
● Total Orders ● Total revenue



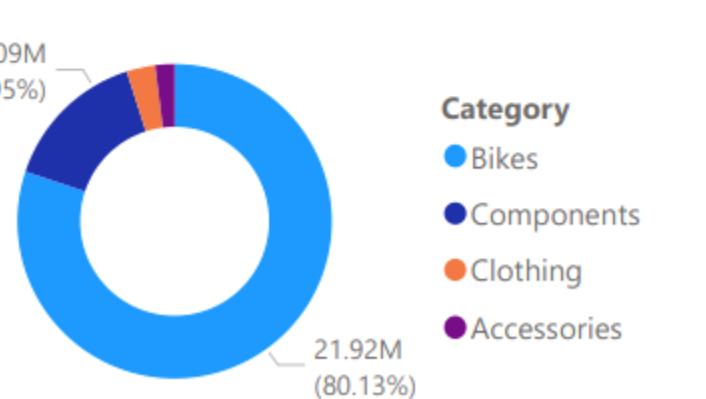
Where are we most present ?



Which subcategory is best sold ?



What product category is selling the most ?



What is our top product by sales

Mountain-200 Black, 38

What is our top product by Total Orders?

Water Bottle - 30 oz

Customer Analysis

AdventureWorks cycles

Product category: All

Calendar: All

Region: All

Profits: 69.51M

Total Sales: 109.81M

Total Orders: 121K

How many orders do we get and how does it impact revenue?

Total Orders (Blue bars) and Total Revenue (Purple line) from January to December 2019.

What is our customers income level?

Donut chart showing IncomeLevel distribution:

IncomeLevel	Value	Percentage
Low	15K	(24.52%)
High	15K	(25.52%)
Average	15K	(24.87%)
Very High	15K	(25.09%)

Which gender is most represented among our customer base?

Donut chart showing Gender distribution:

Gender	Value	Percentage
F	30K	(49.6...)
M	30K	(50.3...)

Who are our top customers?

Jordan Turner: 20 Total Orders, 15,999.10 Total revenue

Willie Xu: 9 Total Orders, 13,490.06 Total revenue

Nichole Nara: 13 Total Orders, 13,295.38 Total revenue

What age group is more represented among our customers?

Stacked bar chart showing Age Group distribution:

Age Group	Count
19	51
20	28
21	58
22	53
23	22
24	24
25	29
26	60
27	41
28	49
29	57
30	23
31	31
32	44
33	40
34	36
35	35
36	36
37	44
38	45
39	48
40	40
41	44
42	47
43	43
44	56
45	50
46	59
47	50
48	42
49	55
50	54
51	56
52	26
53	39

What is our most important channel of distribution?

Total Orders by Channel:

Channel	Total Orders
Reseller	60K
Internet	60K

Type here to search

O File C:/Users/ASUS/AppData/Local/Temp/Power%20BI%20Desktop/print-job-1c112dc6-e0e2-4ea0-a8f3-85692fed629d/Advent...

Page view Read aloud Add text Draw Highlight Erase

ABP SC

EN 27°C 10:50 AM 3/28/2022

Reseller's Analysis

Calendar: 2019 | Region: Germany | Total Sales: 2.27M | Total Costs: 1.89M | Profits: 1.28M | Total Orders: 3704

In which city do we make the most profit?

City	Total revenue
Bad Soden	~320K
Offenbach	~310K
Berlin	~240K
Frankfurt am Main	~230K
Grevenbroich	~200K
München	~170K
Münster	~130K
Saarbrücken	~110K
Eilenburg	~100K
Kiel	~60K
Essen	~40K
Hannover	~30K

Count of Reseller by State-Province

Who are our top resellers?

Warehouse	Total revenue
517	1,258,050.83
Total Orders	

Value Added Reseller	Total revenue
347	576,458.96
Total Orders	

Who is our main reseller?
A Bicycle Association
How much does he make the company?
79.39K

Type here to search

File | C:/Users/ASUS/AppData/Local/Temp/Power%20BI%20Desktop/print-job-1c112dc6-e0e2-4ea0-a8f3-85692fed629d/Adventure... | Page view | Read aloud | Add text | Draw | Highlight | Erase | Print | Back | Forward | Home | Search | Favorites | ABP | SC | Settings | Help | User

3 of 5 | 10:50 AM 3/28/2022 | EN | 27°C | 11

Product Analysis

AdventureWorks cycles

Which product category sells most?

Category	Total Sales
Bikes	~10M
Components	~1M
Clothing	~100K
Accessories	~100K

How much profit do we make over the year?

Month	Total Profits
July	~800K
August	~1.2M
September	~1.5M
October	~300K
November	~3.2M
December	~1.2M

Calendar: 2017 | Region: All | Product category: All

Which products do customers order most?

Product	Order Count	Percentage
Road-150...	254	(23.63%)
Road-150...	252	(23.44%)
Road-150...	204	(18.98%)
Road-150...	183	(16.93%)
...	182	(16.93%)

Which products do customers order least?

Product	Order Count	Percentage
HL Mount...	8	(29.63%)
HL Mount...	7	(25.93%)
HL Road F...	7	(25.93%)
LL Road Fr...	4	(14.81%)
LL Road Fr...	7	(25.93%)

How are we doing compare to our target in terms of orders?

0K 5339 11K

How much profit did we generate compared to our estimates?

0.00M 7.28M 14.55M

① File | C:/Users/ASUS/AppData/Local/Temp/Power%20BI%20Desktop/print-job-1c112dc6-e0e2-4ea0-a8f3-85692fed629d/Advent... ⊖

5 of 5 Q



Calendar

Region

Total Sales

109.81M

Total Costs

97.26M

Profits

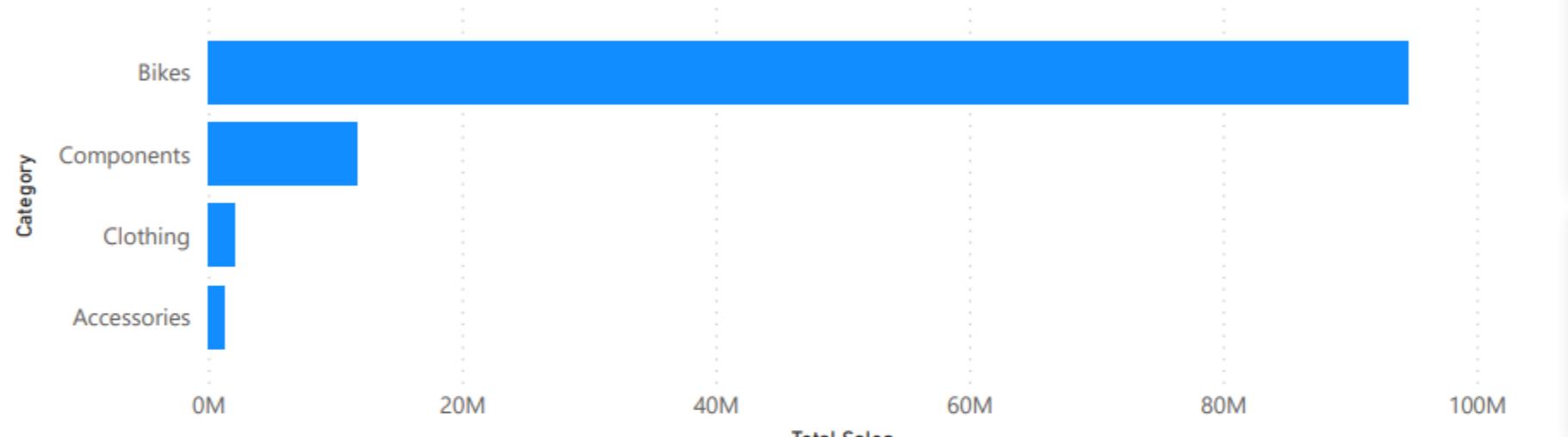
69.51M

Total Orders

121K

Which category of product do we sell most ?

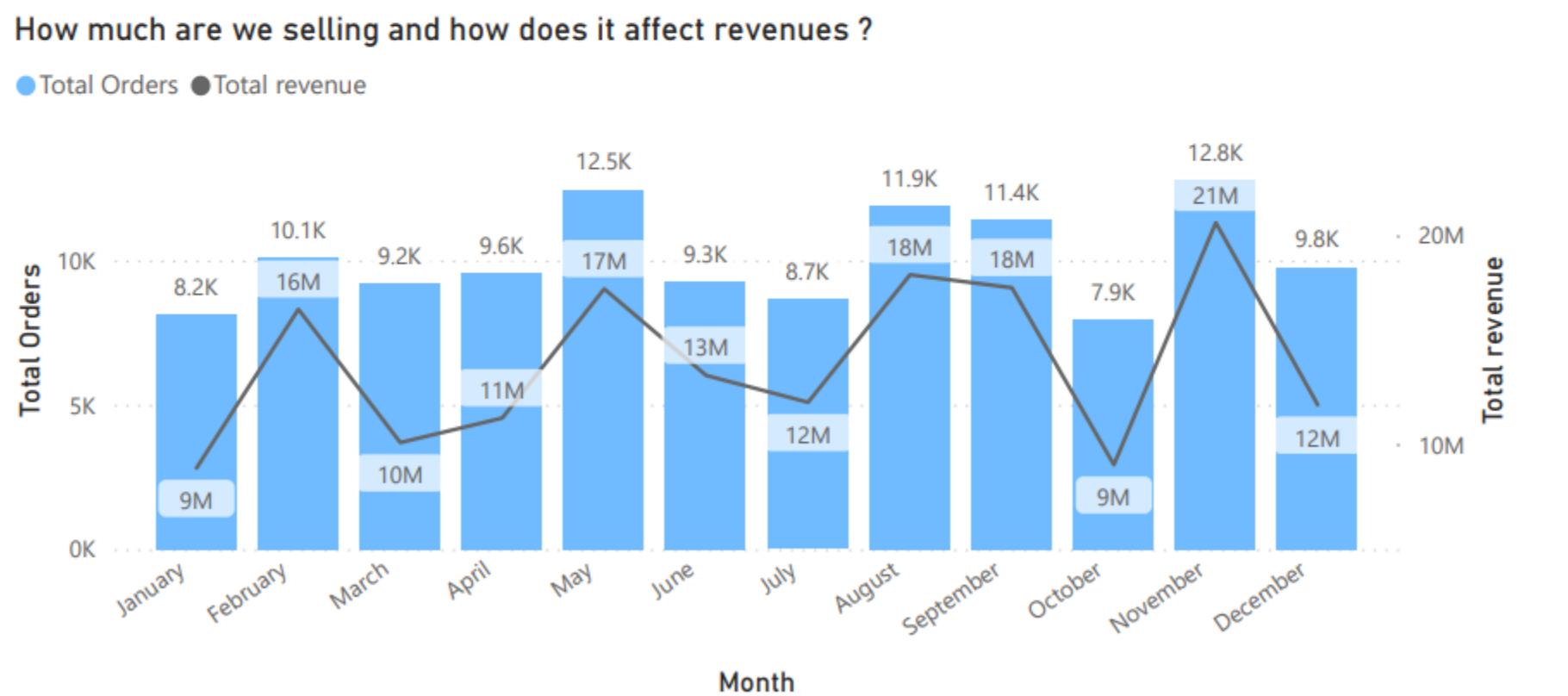
Category



Category	Total Sales
Bikes	~95M
Components	~10M
Clothing	~2M
Accessories	~1M

How much are we selling and how does it affect revenues ?

● Total Orders ● Total revenue



Month	Total Orders	Total Revenue
January	8.2K	9M
February	10.1K	16M
March	9.2K	10M
April	9.6K	11M
May	12.5K	17M
June	9.3K	13M
July	8.7K	12M
August	11.9K	18M
September	11.4K	18M
October	7.9K	9M
November	12.8K	21M
December	9.8K	12M

What product generates the most profit?

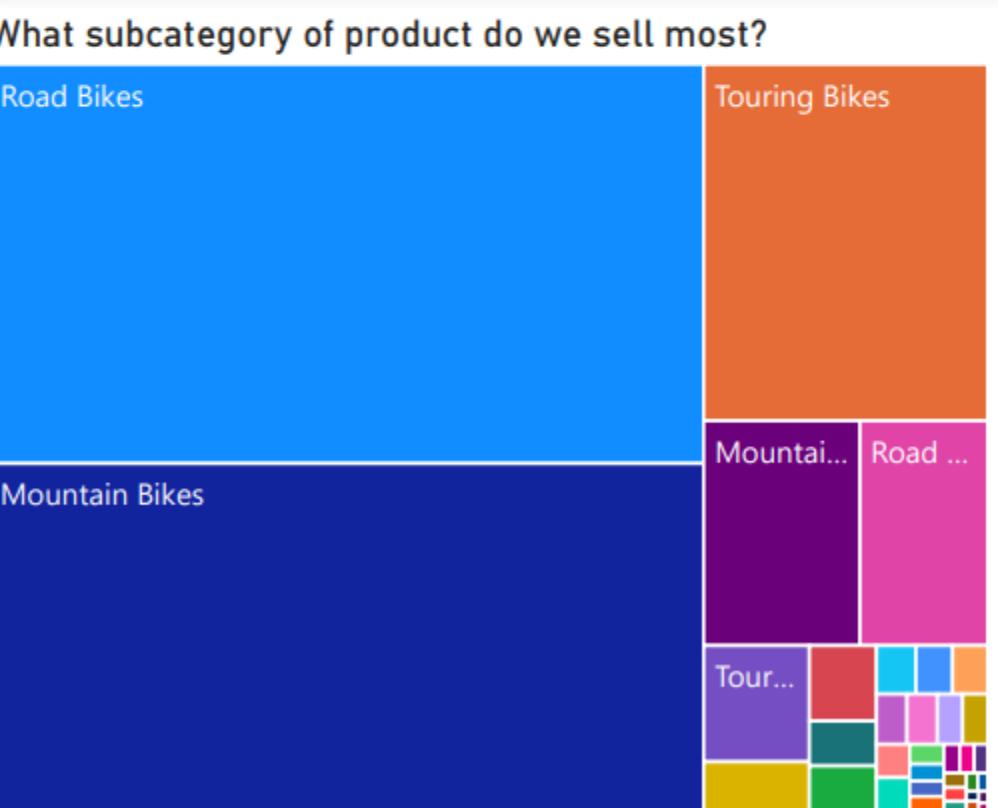
Mountain-200 Black, 38

What is our top seller?

Water Bottle - 30 oz.

What subcategory of product do we sell most?

Road Bikes



Type	Total Sales
Road Bikes	~65M
Mountain Bikes	~25M
Touring Bikes	~15M
Others	~10M

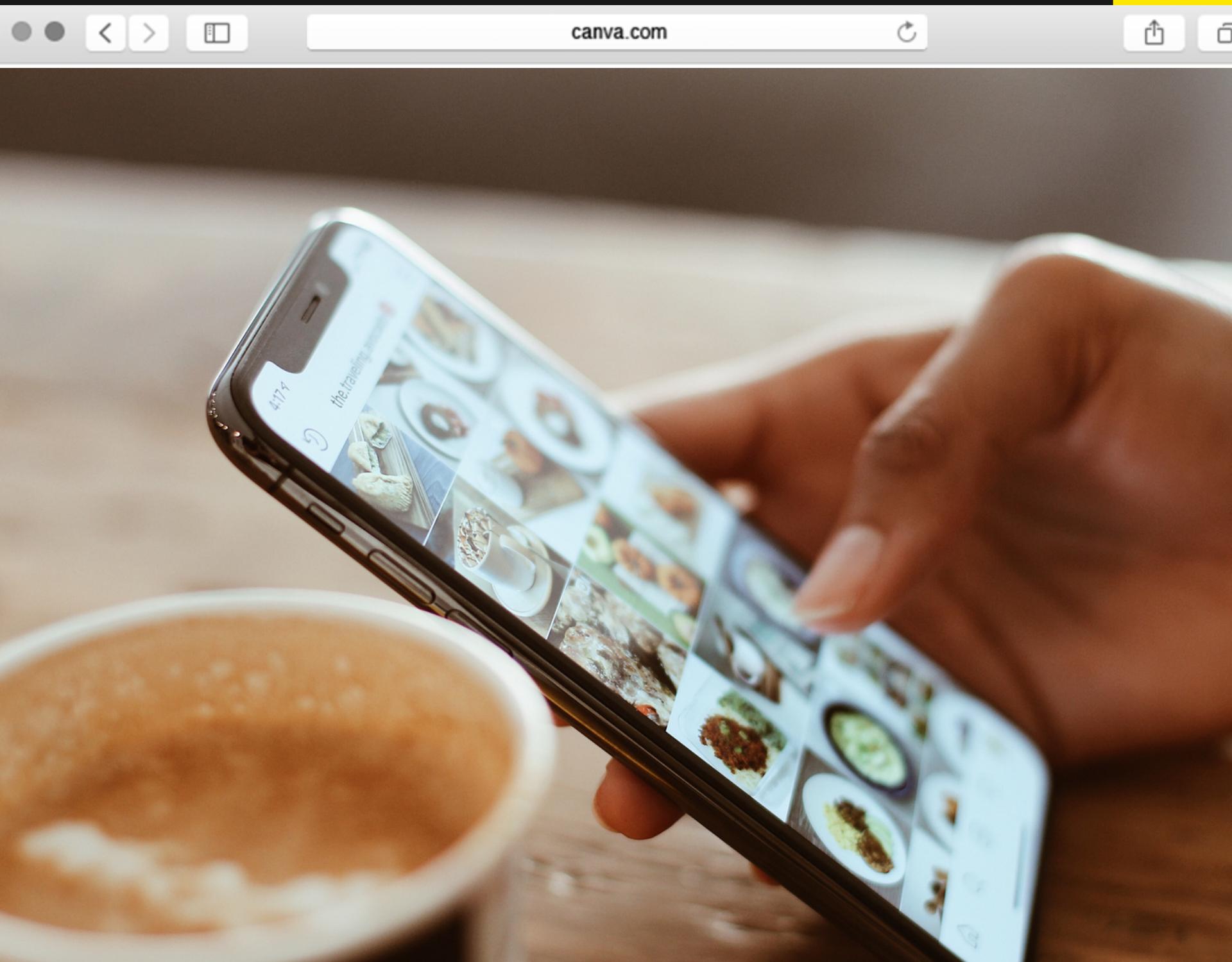
Touring Bikes

Mountain Bikes

Road ...

Mountai...

Tour...



This is the end

We would be glad to answer
any questions

gngoye@edu.hse.ru