

## ASSIGNMENT 2 FRONT SHEET

<b>Qualification</b>	<b>BTEC Level 5 HND Diploma in Computing</b>		
<b>Unit number and title</b>	Unit 14: Business Intelligence		
<b>Submission date</b>	08/01/2022	<b>Date Received 1st submission</b>	08/01/2022
<b>Re-submission Date</b>		<b>Date Received 2nd submission</b>	
<b>Student Name</b>	Le Minh Thien	<b>Student ID</b>	GCH190355
<b>Class</b>	GCH0803	<b>Assessor name</b>	Doan Trung Tung
<b>Student declaration</b>  I certify that the assignment submission is entirely my own work and I fully understand the consequences of plagiarism. I understand that making a false declaration is a form of malpractice.			
		<b>Student's signature</b>	

### Grading grid

P3	P4	P5	P6	M3	M4	D3	D4

⚙ **Summative Feedback:**

⚙ **Resubmission Feedback:**

**Grade:**

**Assessor Signature:**

**Date:**

**IV Signature:**



# **BUSINESS INTELLIGENCE**

Class: GCH0803

# GROUP 2

Trần Quang Huy

Nguyễn Đình Cường

Ngô Tùng Dương

Lê Minh Thiện

Nguyễn Mạnh  
Cường

# Business intelligence

Business intelligence (BI) leverages software and services to transform data into actionable insights that inform an organization's strategic and tactical business decisions. BI tools access and analyze data sets and present analytical findings in reports, summaries, dashboards, graphs, charts and maps to provide users with detailed intelligence about the state of the business.



# Real examples 1: BI helps increase the efficiency of market spending for the company

- **Company:** Univision
- **Challenge:** Univision wanted more visibility into its data to unify and focus on targeted ad campaigns.
- **Solution:** Programmatic TV is an automated and data-driven approach to buying and delivering ads against video content on television, including ads served across the web, mobile devices and connected TVs, as well as linear TV ads served across set-top boxes. With BI powered with information from applications like Facebook, Google Analytics and Adobe Analytics, the company can obtain more value from its programmatic advertising.
- **Results:** Univision achieved an 80% growth in yield during the first quarter after implementing business intelligence.

# Real examples 2: BI to connect departments, data, and processes

- **Company: Stitch Fix**
- **Challenge:** The company wants to reduce returns, keep repeat customers and generate word-of-mouth business with recommendations from customers to their friends and family.
- **Solution:** Stitch Fix collects data within BI throughout the buying process, meaning the more a customer shops with Stitch Fix, the better the styling team comprehends their taste in clothing. The company hired astrophysicists to decode the different personal style components—intricate work that would be impossible without the powerful analytics of BI.
- **Results:** Using business intelligence to profile buyers and their preferences, the company, which started in 2011, reported a customer base of 3.4 million in 2020 and revenues of \$1.7 billion in fiscal year 2020.

# Real examples 3:

## Cementos Argos: BI Improves Financial Efficiency

1. Cementos Argos is a cement company with operations in the U.S., Central and South America and the Caribbean.

1. **Challenge:** The company looked for an overall competitive advantage and a way to support better decision-making.
2. **Solution:** Cementos Argos created a dedicated business analytics center. The company invested in experienced business analysts and data science teams and used BI to leverage data.
3. **Results:** The company standardized the finance process and applied big data to gain more in-depth insight into customer behavior which yielded a higher profitability level.



# BI techniques

## 1.Cleansing

Data cleaning is also known as the process of changing or removing incorrect, corrupted, duplicate or incomplete data in the database. Algorithms and results are unreliable if the data is not correct, even though it appears to be correct. That process is not only interested in deleting data to increase memory for new data, but also finding a way to optimize the accuracy of the data set without having to delete the information. It removes data including spelling, modifies missing code snippets, syntax errors, identifies duplicates. Data cleaning plays a big role in developing reliable answers and in analytics which is considered a fundamental feature of data science. The engine of the data cleaning service is to build a uniform and standardized data set for business intelligence and data analytics tools to easily capture and know the exact perceptions of problems.



## **Labeling**

Data labeling is the process of tagging and detecting data patterns. This process can be done manually but is usually done by software assistance.

## 2. Analysis techniques

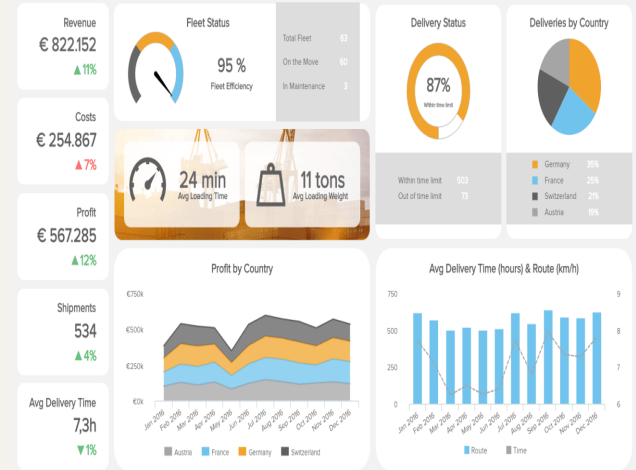
### Queries

The query invokes information from a database table or a data request or combination of tables. This type of data can be produced as results returned by the structured query language (SQL) or as graphs, complex results, or as images. Among the various query languages that can be used to perform database queries from the most complex to the simplest. SQL, is a famous and most commonly used query language and it is familiar to most database administrators today.

## 2. Analysis techniques

### Dashboards

Dashboard is a BI software interface that provides metrics, statistics, insights, and visualizations pre-configured or customer-added to existing data. It allows end users and power users alike to see immediate, immediate results of business or data analytics operations. A component of every BI software solution, the BI dashboard acts as a standard application dashboard that brings together two or more key performance indicators (KPIs) and an intuitive interface statistics page. When the user uses the software the type and amount of data displayed varies according to the capabilities and customizations available to the software.



# Analytic techniques

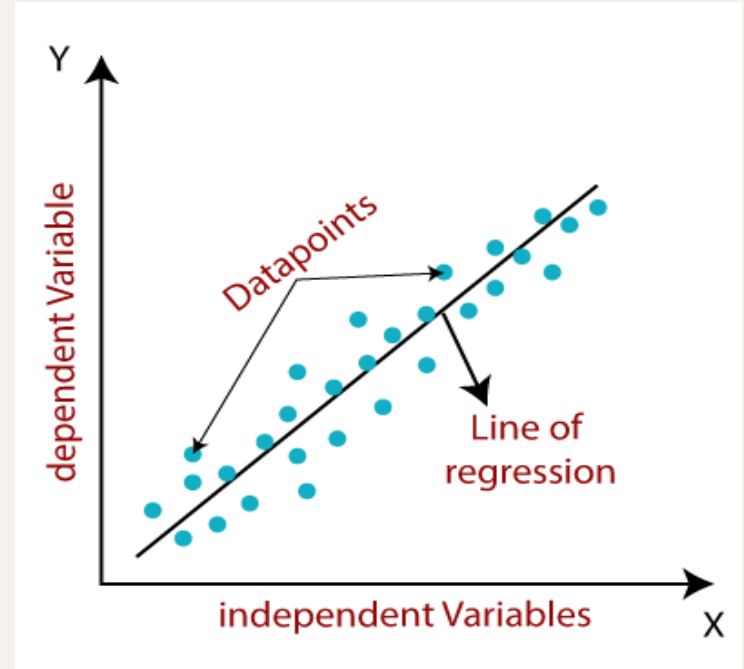
## Regression

Regression analysis is a reliable method to use to determine which variables might influence a topic of interest. Real-time regression allows the user to determine exactly which factors are most important, which factors can be ignored, and how these factors affect each other.

To understand regression analysis, there are two terms that must be understood:

**Dependent variable:** is the main factor that the user tries to understand or predict.

**Independent variables:** are the factors that the user must hypothesize have an impact on the user's dependent variable.



# Machine learning

## Machine learning

Machine learning is an application of artificial intelligence (AI) that gives systems the ability to automatically learn and improve from experience, without the need for humans to explicitly program them. Machine learning focuses on developing computer programs that can access data and use it for self-learning. The machine learning process starts with observations or data. The main purpose is to allow computers to learn automatically without human intervention and adjust operations accordingly.



# BI Tool

# Tableau

Tableau is a data visualization tool with a user-friendly interface, easy to use basic features. When designing the Dashboard, you need to create a worksheet and drag it into the container.



## Business Intelligence feature:

- Supports multiple data sources: Relational database, NoSQL database, Multi-dimensional database, Big Data Platform, File data sources (Excel, csv, txt, Json, pdf, mdb, Tableau)
- Powerful data visualization capabilities.
- Using unique visualization technology, Tableau can quickly analyze data by presenting analysis results in color, shape, and size. Tableau's development team continues to work on the types of charts that users love. Therefore, Tableau can best display the graph of the data. However, gauges, 3D graphs and heat maps cannot be used.



# QlikView

QlikView is a self-service tool that works with its own association technology and internal memory technology. Intuitive interface, easy to use, comes with smart search feature, no need to build blocks, suitable for academic analysis rather than daily analysis.



## Business Intelligence feature

- In-memory technology: QlikView computes data while it moves from backend to frontend. Since no aggregated or pre-computed data is stored, system memory can be saved and data transfer speeds are fast.
- Search for direct and indirect data: You can find both direct and indirect data. Directly, you can input data and get information, indirectly, input the desired data related content and get all related data.

# Sisense

Sisense has an intuitive interface, operating according to the drag and drop mechanism. The online community is large, although not as crowded as Tableau, but the resources here are enough for newcomers to get acquainted and use.



## Business Intelligence feature

- Fast Speed: Ask any question and get instant answers without going back to the drawing board for new queries thanks to the engine software in Chip.
- Flexible Integration: Easily integrates with third-party applications such as Google Adwords, Excel, Zendesk, and Salesforce.

# Dataset

The dataset contains historical car auction sales prices, scraped from the outside internet sources. The dataset has been collected in 2015, and will not be updated.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	year	make	model	trim	body	transmission	vin	state	condition	odometer	color	interior	seller	mmr	sellingprice	saledate			
2	2015	Kia	Sorento	LX	SUV	automatic	5xyktca69fg566472	ca	5	16639	white	black	kia motors america; inc	20500	21500	Tue Dec 16 2014 12:30:00 GMT-0800 (PST)			
3	2015	Kia	Sorento	LX	SUV	automatic	5xyktca69fg561319	ca	5	9393	white	beige	kia motors america; inc	20800	21500	Tue Dec 16 2014 12:30:00 GMT-0800 (PST)			
4	2014	BMW	3 Series	328i SULEV	Sedan	automatic	wba3c1c51ek116351	ca	4.5	1331	gray	black	financial services remark	31900	30000	Thu Jan 15 2015 04:30:00 GMT-0800 (PST)			
5	2015	Volvo	S60	T5	Sedan	automatic	yy1612tb4f13110987	ca	4.1	14282	white	black	volvo na rep/world omn	27500	27750	Thu Jan 29 2015 04:30:00 GMT-0800 (PST)			
6	2014	BMW	6 Series Gr	650i	Sedan	automatic	wba6b2c57ed129731	ca	4.3	2641	gray	black	financial services remark	66000	67000	Thu Dec 18 2014 12:30:00 GMT-0800 (PST)			
7	2015	Nissan	Altima	2.5 S	Sedan	automatic	1n4al3ap1fn326013	ca	1	5554	gray	black	enterprise vehicle excha	15350	10900	Tue Dec 30 2014 12:00:00 GMT-0800 (PST)			
8	2014	BMW	M5	Base	Sedan	automatic	wbsfv9c51ed593089	ca	3.4	14943	black	black	the hertz corporation	69000	65000	Wed Dec 17 2014 12:30:00 GMT-0800 (PST)			
9	2014	Chevrolet	Cruze	1LT	Sedan	automatic	1g1pc5sb2e7128460	ca	2	28617	black	black	enterprise vehicle excha	11900	9800	Tue Dec 16 2014 13:00:00 GMT-0800 (PST)			
10	2014	Audi	A4	2.0T Premium Plus	Sedan	automatic	wauffaf13en030343	ca	4.2	9557	white	black	audi mission viejo	32100	32250	Thu Dec 18 2014 12:00:00 GMT-0800 (PST)			
11	2014	Chevrolet	Camaro	LT	Convertible	automatic	2g1fb3d37e9218789	ca	3	4809	red	black	d/m auto sales inc	26300	17500	Tue Jan 20 2015 04:00:00 GMT-0800 (PST)			
12	2014	Audi	A6	3.0T Prestige quattro	Sedan	automatic	wauhgafcoen062916	ca	4.8	14414	black	black	desert auto trade	47300	49750	Tue Dec 16 2014 12:30:00 GMT-0800 (PST)			
13	2015	Kia	Optima	LX	Sedan	automatic	5xxgm4a73fg353538	ca	4.8	2034	red	tan	kia motors finance	15150	17700	Tue Dec 16 2014 12:00:00 GMT-0800 (PST)			
14	2015	Ford	Fusion	SE	Sedan	automatic	3fa6p0hdxr145753	ca	2	5559	white	beige	enterprise vehicle excha	15350	12000	Tue Jan 13 2015 12:00:00 GMT-0800 (PST)			
15	2015	Kia	Sorento	LX	SUV	automatic	5xyktca66fg561407	ca	5	14634	silver	black	kia motors america; inc	20600	21500	Tue Dec 16 2014 12:30:00 GMT-0800 (PST)			
16	2014	Chevrolet	Cruze	2LT	Sedan	automatic	1g1pe5sbxe7120097	ca		15686	blue	black	avis rac/san leandro	13900	10600	Tue Dec 16 2014 12:00:00 GMT-0800 (PST)			
17	2015	Nissan	Altima	2.5 S	Sedan	automatic	1n4al3ap5fc124223	ca	2	11398	black	black	enterprise vehicle excha	14750	14100	Tue Dec 23 2014 12:00:00 GMT-0800 (PST)			
18	2015	Hyundai	Sonata	SE	Sedan		5npe24af4fh001562	ca		8311	red	â€”	avis tra	15200	4200	Tue Dec 16 2014 13:00:00 GMT-0800 (PST)			
19	2014	Audi	Q5	2.0T Premium Plus	SUV	automatic	wa1fafpxea085074	ca	4.9	7983	white	black	audi north scottsdale	37100	40000	Thu Dec 18 2014 12:30:00 GMT-0800 (PST)			
20	2014	Chevrolet	Camaro	LS	Coupe	automatic	2g1fa1e39e9134494	ca	1.7	13441	black	black	wells fargo dealer servic	17750	17000	Tue Dec 30 2014 15:00:00 GMT-0800 (PST)			
21	2014	BMW	6 Series	650i	Convertible	automatic	wbayp9c53ed169260	ca	3.4	8819	black	black	the hertz corporation	68000	67200	Wed Dec 17 2014 12:30:00 GMT-0800 (PST)			
22	2015	Chevrolet	Impala	LTZ	Sedan	automatic	2g1165s30f9103921	ca	1.9	14538	silver	black	enterprise vehicle excha	24300	7200	Tue Jul 07 2015 09:30:00 GMT-0700 (PDT)			
23	2014	BMW	5 Series	528i	Sedan	automatic	wba5a5c51ed501631	ca	2.9	25969	black	black	financial services remark	34200	30000	Tue Feb 03 2015 04:30:00 GMT-0800 (PST)			
24	2014	Chevrolet	Camaro	LT	Convertible	automatic	2g1fb3d31e9134662	ca		33450	black	black	avis rac/san leandro	20100	14700	Tue Dec 16 2014 12:00:00 GMT-0800 (PST)			
25	2015	Audi	A3	1.8 TFSI Premium	Sedan	automatic	wauacgff7f1002327	ca	4.9	5826	gray	black	audi north scottsdale	24000	23750	Thu Dec 18 2014 12:30:00 GMT-0800 (PST)			
26	2014	BMW	6 Series	650i	Convertible	automatic	wbayp9c57ed169262	ca	3.8	10736	black	black	the hertz corporation	67000	65000	Tue Jan 06 2015 12:30:00 GMT-0800 (PST)			
27	2015	Hyundai	Sonata	SE	Sedan	automatic	5npe24af4fh038482	ca		9281	silver	gray	enterprise vehicle excha	15150	8500	Tue Dec 16 2014 13:00:00 GMT-0800 (PST)			
28	2015	Volvo	XC70	T6	Wagon	automatic	yy4902nb3f1198103	ca	4.2	16506	brown	brown	volvo na rep/world omn	32100	32500	Thu Feb 26 2015 04:30:00 GMT-0800 (PST)			
29	2015	Volvo	XC70	T6	Wagon	automatic	yy4902nb3f1196951	ca	4.8	12725	beige	beige	volvo na rep/world omn	32300	32500	Thu Feb 12 2015 04:30:00 GMT-0800 (PST)			
30	2014	BMW	5 Series	528i	Sedan	automatic	wba5a5c51ed501631	ca		14538	silver	black	financial services remark	34200	30000	Tue Feb 03 2015 04:30:00 GMT-0800 (PST)			

---

# Dataset Requirements

-The data has been converted to semi-structured data, but there are still many places that are not clear and complete for analysis. So we decided to process the data for better analysis. Below are some of the requirements needed to complete the data.

- Delete column vin.
  - Merge the model and trim columns.
  - In the transmission column, change null to nonautomatic.
  - Make a dictionary of the list of states and convert them to full names.
  - At the odometer column, round to hundreds.
  - The value of condition will multiply by 10.
  - Time data will drop the time keeping only the day, month, and year.
-

# Read file process

The above code is used to read data from the dataset file that we have prepared, we have read CSV files with Python built-in CSV library

```
1  ### open a csv file
2  import csv # import csv library
3  f = open('car.csv') # open file car.csv
4  cars = [] # create a cars list
5  for row in csv.reader(f, delimiter=','): #read lines in csv file
6      cars.append(row) # append the lines to cars list
7
8  f.close() # close file
9
```

# Data processing

```
1  ### delete an column
2  for row in cars[:]:
3      row.remove(row[6])
4
```

This block of code is used to remove a column from the data list.

This block of code below is used to merge two columns of data together.

```
1  ### Match columns model and trim
2  for row in cars[:]:
3      row[2] = row[2] + ' ' + row[3] # modify column 2 = column 2 + column 3
4      row.remove(row[3]) # delete column 3
```

# Data processing

The above code is used to round the data for the odometer column. Data will be rounded to the hundredth.

```
1  ### accomplish odometer
2  for row in cars[1:]:
3      if row[7]:
4          i = (float(row[7]) // 100)
5          if i != 0:
6              if float(row[7]) % 100 >= 50:
7                  row[7] = (i + 1) * 100
8              else:
9                  row[7] = i * 100
10
11
```

*Comments:*

- #if value not null or empty string => do statements
- #find the integer part divisible by 100
- #If value >= 100 => do statements
- #if remainder of divisible by 100 >=50 => do statements
- #New value = the integer part divisible by 100 add 1 and core 100
- #if remainder of divisible by 100 <= 50 =< do statements
- #New value = the integer part divisible by 100 and core 100

# Data processing

The above code block is used to filter the date data and keep only the day, month and year.

```
1  ### Filter the time data and keep only the date, month and year
2  for row in cars[1:]:
3      a = []
4      a = row[13].split(" ")[1:4] # Split the date, month and year in time data
5      b = ' '.join(a) # Convert the list of date, month and year to a string
6      row.remove(row[13]) # Remove the old datetime variable
7      row.append(b) # add the new datetime variable
8
```



# Data processing

The code block below is used to read data from a state library file, the loop will check if the state abbreviation of the dataset defined in the library or not and replace it with the full name.

```
1 ### Add state library
2 import csv
3 m = open('state.csv')    # open file state.csv
4 states = []
5 for row in csv.reader(m, delimiter=','):
6     states.append(row)
7 m.close()

1 ### Replace state name
2 ac = [row[1] for row in states] # Create a list of states abbreviation
3
4 for row in cars[1:]:
5     if ac.count(row[5].upper()): # if data exist in state library => do statements
6         row[5] = ''.join(states[ac.index(row[5].upper())[:-1]]) #Replace old data by state fullname
```

# Data processing

The code block below is used to multiply the value of condition by 10. We have set Try - except to avoid an error when the data is invalid.

```
1 ### Multiply the value of condition by 10
2 for row in cars[1:]:
3     if row[6]: #if variable not null or empty => do statements
4         try:
5             row[6] = float(row[6]) * 10.0 # Try exception when convert to float
6         except ValueError: # if ValueError (can not convert to float) set value to 0
7             row[6] = 0
```

# Read file process

The above code block is used to output the data to a new csv file. We used the built-in python library to perform the action.

```
1 import csv
2 with open('asm.csv', 'w', newline='') as file: # open asm.csv file in mode write
3     writer = csv.writer(file)                 # Define writer by csv.writer() Function
4     writer.writerow(cars)                     # Save Car's data to csv file
```

# DataSet after Processing

This is the dataset after it has been processed. Columns have become clearer and some unnecessary data has been removed.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	year	make	model trim	body	transmission	state	condition	odometer	color	interior	seller	mmr	sellingprice	saledate
2	2015	Kia	Sorento LX	SUV	automatic	California	50.0	16600.0	white	black	kia motors america; inc	20500	21500	Dec 16 2014
3	2015	Kia	Sorento LX	SUV	automatic	California	50.0	9400.0	white	beige	kia motors america; inc	20800	21500	Dec 16 2014
4	2014	BMW	3 Series 328i SULEV	Sedan	automatic	California	45.0	1300.0	gray	black	financial services remarketing (lease)	31900	30000	Jan 15 2015
5	2015	Volvo	S60 T5	Sedan	automatic	California	41.0	14300.0	white	black	volvo na rep/world omni	27500	27750	Jan 29 2015
6	2014	BMW	6 Series Gran Coupe 650i	Sedan	automatic	California	43.0	2600.0	gray	black	financial services remarketing (lease)	66000	67000	Dec 18 2014
7	2015	Nissan	Altima 2.5 S	Sedan	automatic	California	10.0	5600.0	gray	black	enterprise vehicle exchange / tra / rental / tulsa	15350	10900	Dec 30 2014
8	2014	BMW	M5 Base	Sedan	automatic	California	34.0	14900.0	black	black	the hertz corporation	69000	65000	Dec 17 2014
9	2014	Chevrolet	Cruze 1LT	Sedan	automatic	California	20.0	28600.0	black	black	enterprise vehicle exchange / tra / rental / tulsa	11900	9800	Dec 16 2014
10	2014	Audi	A4 2.0T Premium Plus quattro	Sedan	automatic	California	42.0	9600.0	white	black	audi mission viejo	32100	32250	Dec 18 2014
11	2014	Chevrolet	Camaro LT	Convertible	automatic	California	30.0	4800.0	red	black	d/m auto sales inc	26300	17500	Jan 20 2015
12	2014	Audi	A6 3.0T Prestige quattro	Sedan	automatic	California	48.0	14400.0	black	black	desert auto trade	47300	49750	Dec 16 2014
13	2015	Kia	Optima LX	Sedan	automatic	California	48.0	2000.0	red	tan	kia motors finance	15150	17700	Dec 16 2014
14	2015	Ford	Fusion SE	Sedan	automatic	California	20.0	5600.0	white	beige	enterprise vehicle exchange / tra / rental / tulsa	15350	12000	Jan 13 2015
15	2015	Kia	Sorento LX	SUV	automatic	California	50.0	14600.0	silver	black	kia motors america; inc	20600	21500	Dec 16 2014
16	2014	Chevrolet	Cruze 2LT	Sedan	automatic	California		15700.0	blue	black	avis rac/san leandro	13900	10600	Dec 16 2014
17	2015	Nissan	Altima 2.5 S	Sedan	automatic	California	20.0	11400.0	black	black	enterprise vehicle exchange / tra / rental / tulsa	14750	14100	Dec 23 2014
18	2015	Hyundai	Sonata SE	Sedan	nonautomatic	California		8300.0	red	â€	avis tra	15200	4200	Dec 16 2014
19	2014	Audi	Q5 2.0T Premium Plus quattro	SUV	automatic	California	49.0	8000.0	white	black	audi north scottsdale	37100	40000	Dec 18 2014
20	2014	Chevrolet	Camaro LS	Coupe	automatic	California	17.0	13400.0	black	black	wells fargo dealer services	17750	17000	Dec 30 2014
21	2014	BMW	6 Series 650i	Convertible	automatic	California	34.0	8800.0	black	black	the hertz corporation	68000	67200	Dec 17 2014
22	2015	Chevrolet	Impala LTZ	Sedan	automatic	California	19.0	14500.0	silver	black	enterprise vehicle exchange / tra / rental / tulsa	24300	7200	Jul 07 2015
23	2014	BMW	5 Series 528i	Sedan	automatic	California	29.0	26000.0	black	black	financial services remarketing (lease)	34200	30000	Feb 03 2015
24	2014	Chevrolet	Camaro LT	Convertible	automatic	California		33500.0	black	black	avis rac/san leandro	20100	14700	Dec 16 2014
25	2015	Audi	A3 1.8 TFSI Premium	Sedan	automatic	California	49.0	5800.0	gray	black	audi north scottsdale	24000	23750	Dec 18 2014
26	2014	BMW	6 Series 650i	Convertible	automatic	California	38.0	10700.0	black	black	the hertz corporation	67000	65000	Jan 06 2015
27	2015	Hyundai	Sonata SE	Sedan	automatic	California		9300.0	silver	gray	enterprise vehicle exchange / tra / rental / tulsa	15150	8500	Dec 16 2014
28	2015	Volvo	XC70 T6	Wagon	automatic	California	42.0	16500.0	brown	brown	volvo na rep/world omni	32100	32500	Feb 26 2015

---



# **Design Dashboards**

---

---

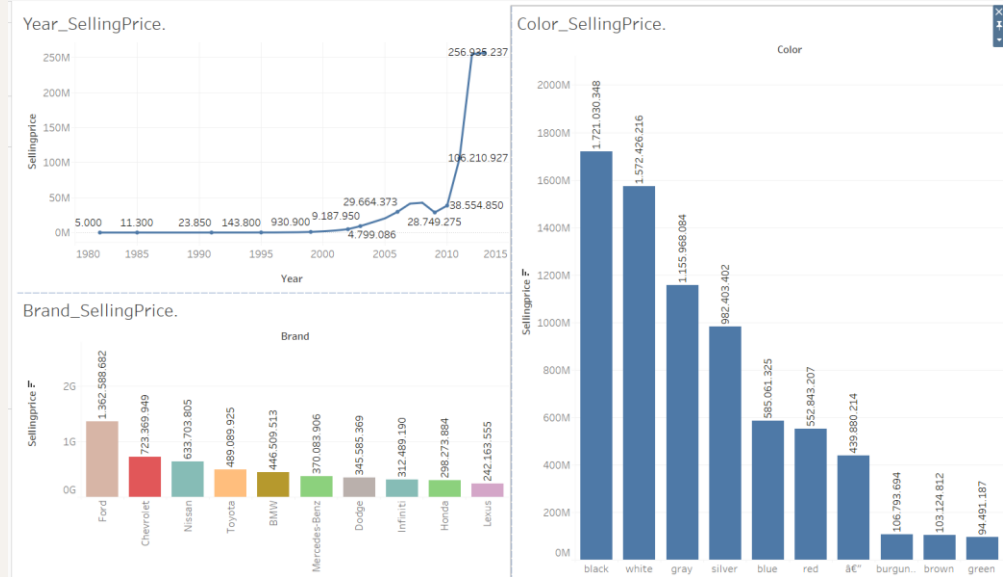
## **Dashboard 1: Dashboard 1: Brand and color trends based on selling price over the years**

# Dashboard 1

## Purpose:

This dashboard includes 3 charts, selling price by year, best selling brand by revenue and preferred color by revenue. It is designed to know what car brand and color trends users like over the years. To have an insight into the trends of the times so that the company can rely on it to adapt to today's trends in brand and color. It's the useful information that dashboards bring to our company's operations and we can make better business decisions.

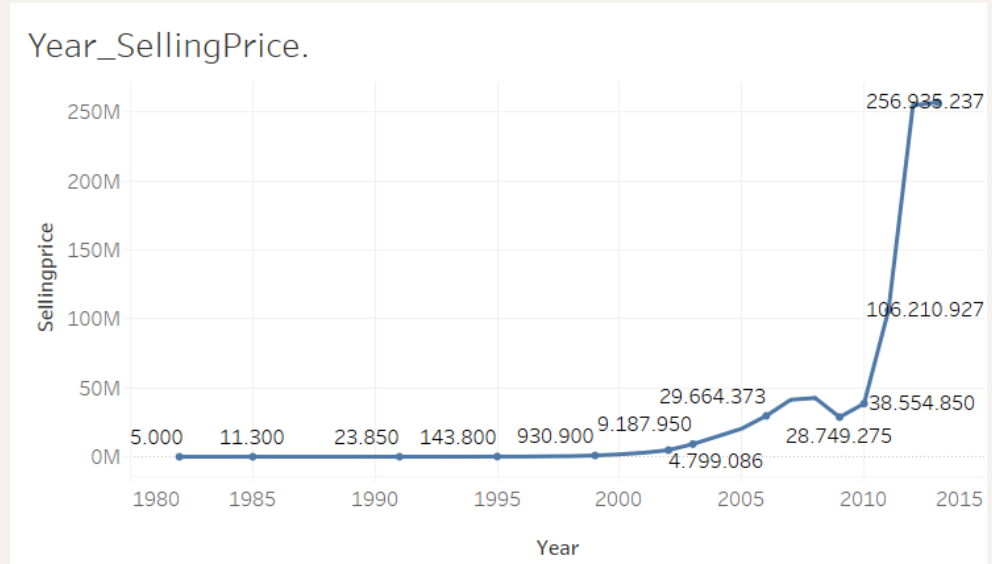
## Illustration:



# Chart 1: Company's sellingprice by year

## Purpose:

This chart is designed as a line chart based on the results of the selling price over the years. As you can see the left column is the sellingprice column and the line below is the years row. Therefore, combined with other variables, we can easily observe the company's selling price over the years

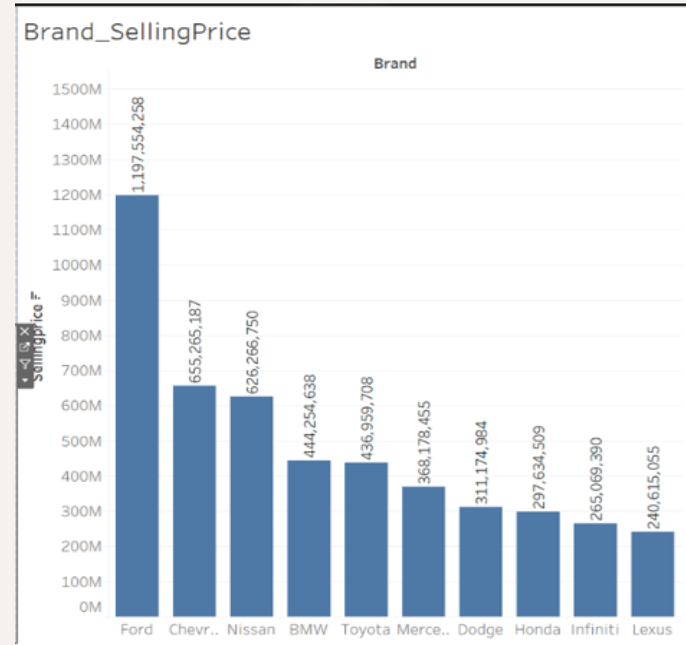




## Chart 2: Car brand trends by selling price

### Purpose:

This chart is designed as a bar chart with the vertical axis being the selling price and the horizontal axis being the car brand. Through the chart, it is easy to see which brand is trending based on the selling price on the top of the column. Thereby, it can help the company see which brands are trending and make decisions in the future.

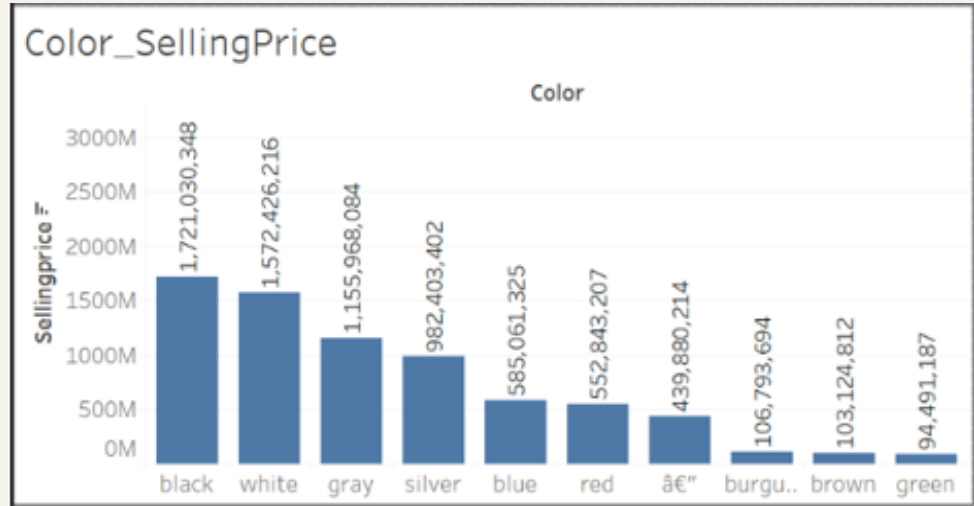


# Chart 3: sellingprice based on trend color

## Purpose:

This chart is designed as a bar chart with the vertical axis being the selling price and the horizontal axis being the car color.

Through the chart, it is easy to see which brands are trending based on the selling prices on the top of the column. Thereby, it can help the company see which colors are trending and make decisions in the future.



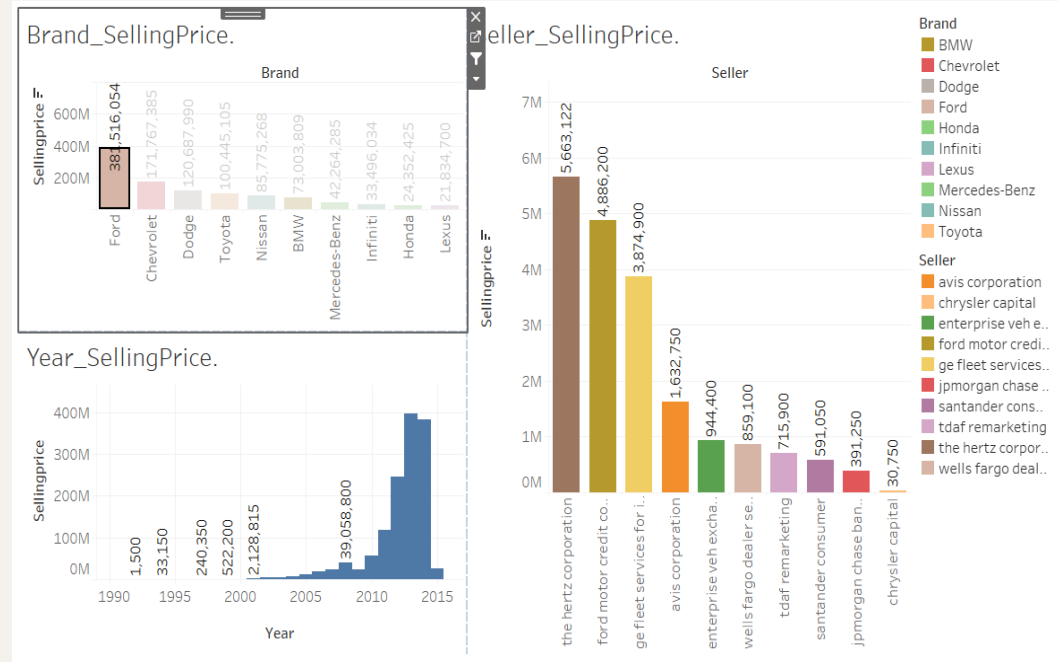
# Interaction on Dashboard 1

**Suggestion:**

Through the conclusions drawn from the interactions on the dashboard, it will help us easily assess which car brands and colors are trending over the years. Specifically, in 2014 with the highest sales (120.687,990), the brand that was the trend of that year was Ford and the most popular color was white. Moreover, the company can rely on the collected interaction data to know what the trends are so that they can make appropriate strategies for the company.

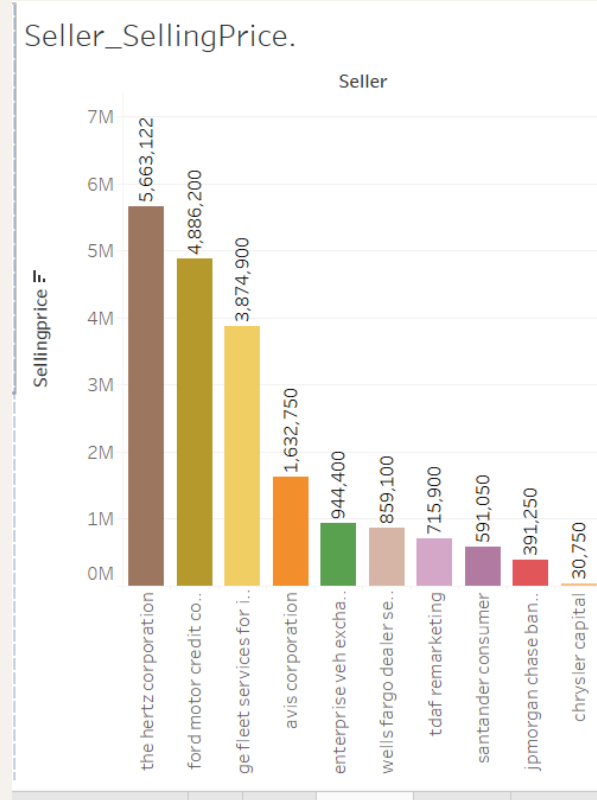
# Dashboard 2

**Purpose:** This dashboard includes 3 charts, selling price by year, best selling brand by revenue and highest auction base by sales (seller). It is designed to know through the brand that it will be possible to know which brand has the highest sales in which year and the Year\_SellingPrice chart will display the data related to the car manufacturer of your choice and also It is possible to know at which base (seller) has the highest auction turnover of the selected car company. The analysis will help the company have an insight into the trend of the times so that the company can rely on it to adapt to today's brand trends. It is useful information that the dashboard brings to our company operations and we can make better business decisions



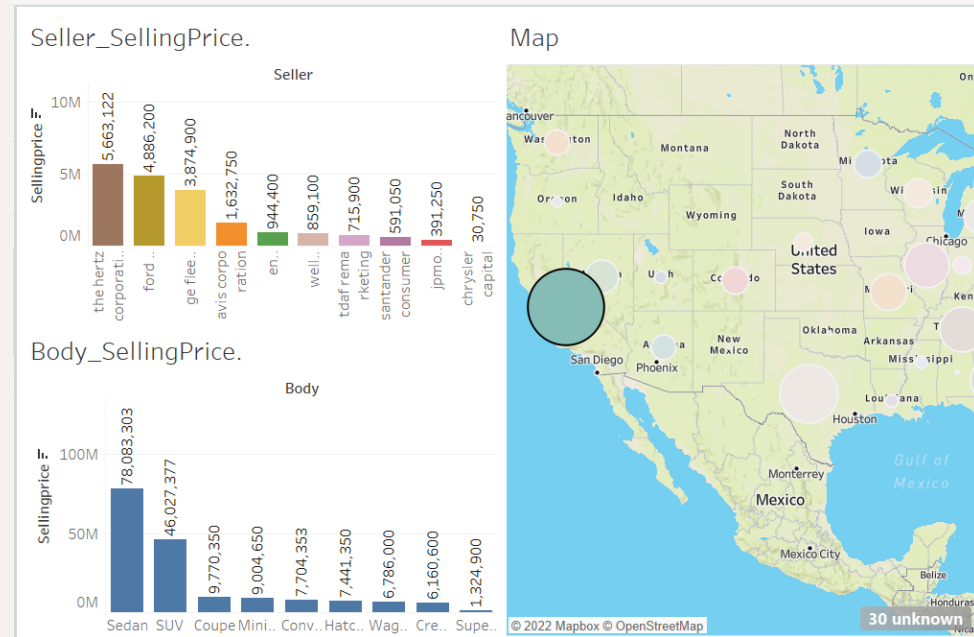
# Chart 1: selling price statistics of sellers

This chart is designed as a column chart based on the results of product selling prices of distribution stores (sellers). You can see the left part is the selling price and the line below is the corresponding seller. We can rely on the chart to know the order of the sellers through the seller's price



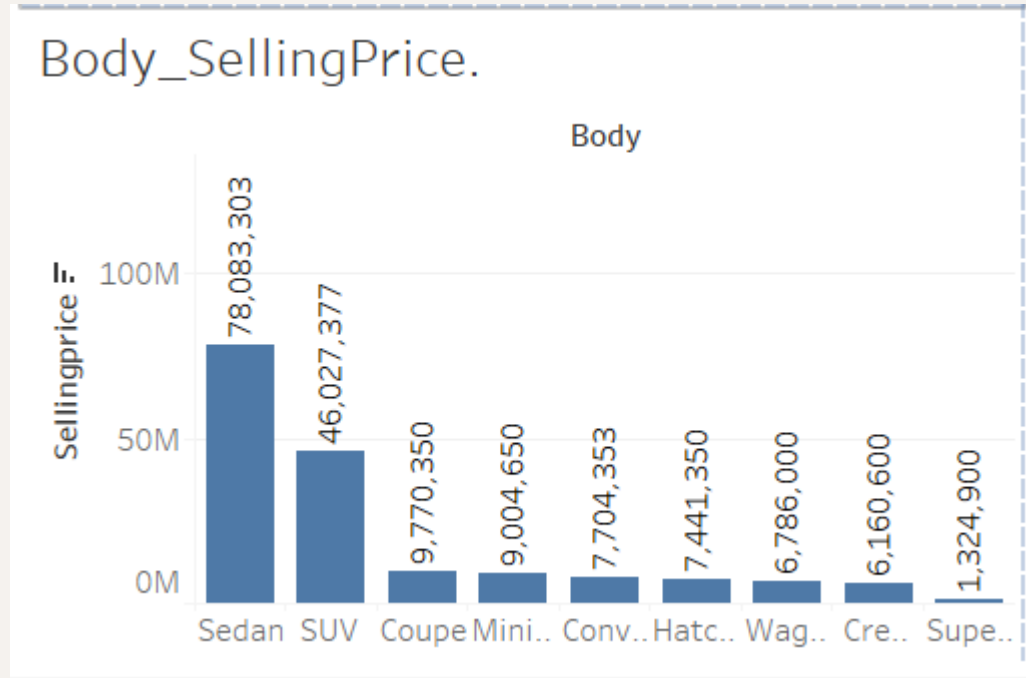
# Dashboard 3

This dashboard includes 3 charts, seller by price, best-selling vehicle by revenue, and a chart showing US states. It is designed to find out which seller has the highest sales and which vehicles are sold the most in the states. To get an in-depth look at the market and vehicle type by state, determine which states have the most auto auctioneers and also know which vehicles are the most popular in the states. It is useful information that dashboards bring to our company operations and we can make better business decisions.

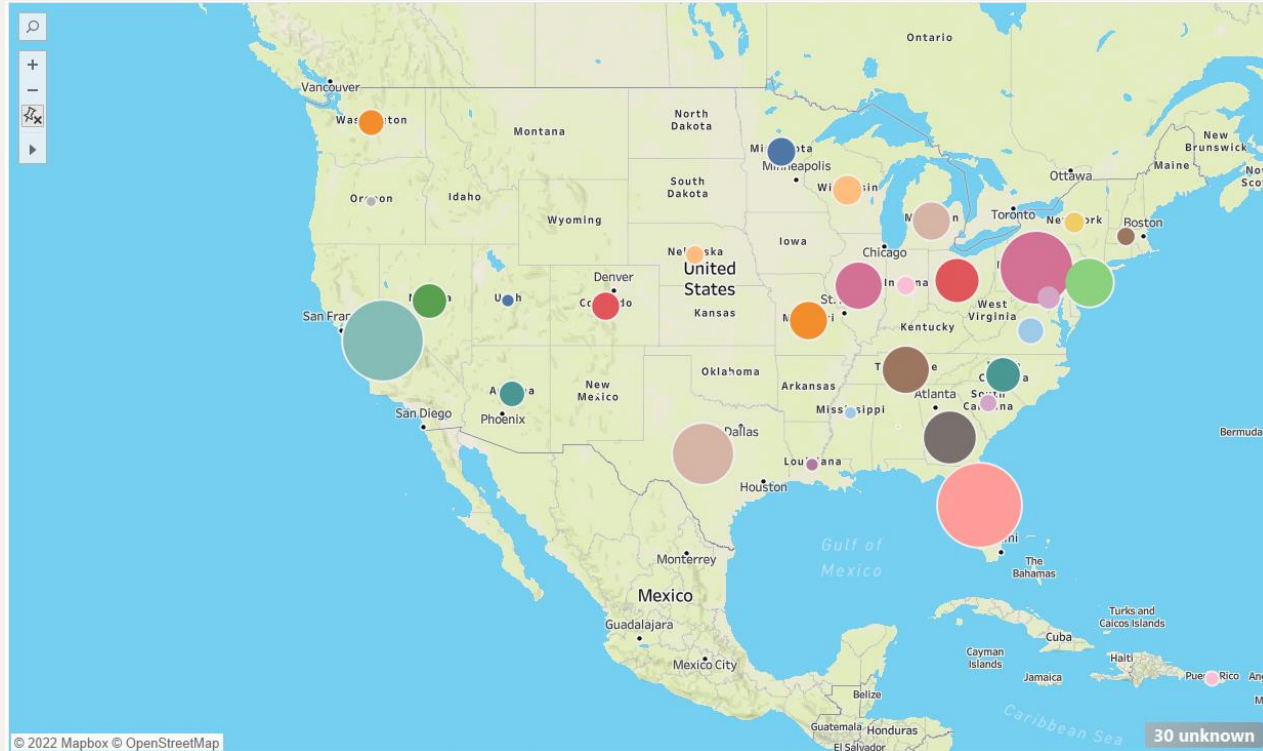


# Chart 1

This chart is designed with vertical column format as selling price and horizontal column as vehicle type. Through the chart, you can clearly see which type of car is being auctioned by the most people. Thereby, it can help the company know all types of vehicles that are trending among users and can make future development decisions.



# Map chart

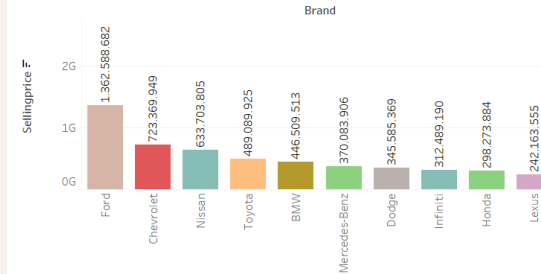




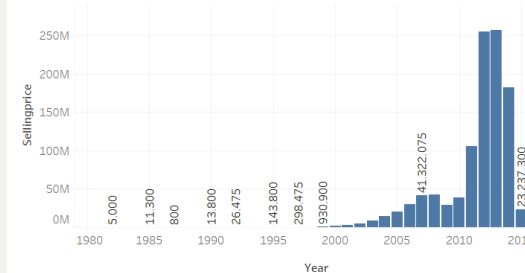
# Dashboard 4

Purpose: This dashboard includes 3 charts, which map selling prices by state, selling price by year, and best-selling brands by revenue. It is designed to know the selling price and brand of the states year by year. To have an insight into the trend of the times so that the company can base on that to adapt to the trend of brands and locations. It is useful information that dashboards bring to our company operations and we can make better business decisions.

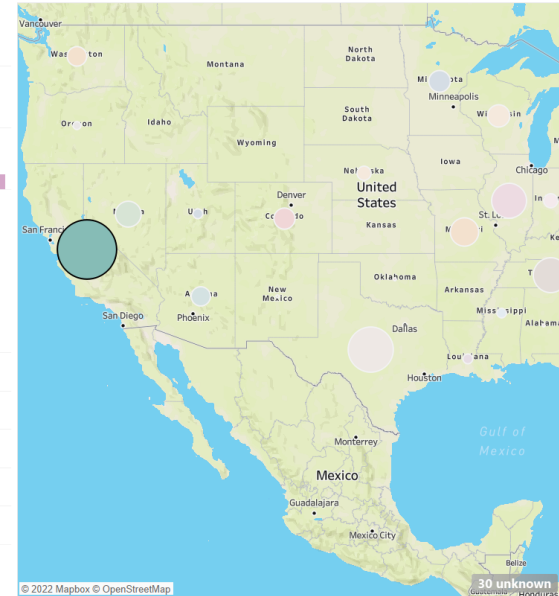
Brand\_SellingPrice.



Year\_SellingPrice.



Map



# Review feed-back

## Good points:

- Data after processing has been cleaned. The data no longer contains nulls or irrelevant values. This makes the data cleaner.
- For code processing. Code blocks are clearly allocated for each request making it easy for the user to visualize how it works.
- The use of techniques of the Numpy library will help a lot in improving .
- The charts and sheets are designed clearly and in detail. Besides, the coloring helps the viewer to easily distinguish the data from each other.
- The dashboard has many useful interactions to help users easily visualize the effects of the charts on each other.
- The content is not duplicated so that users can comfortably learn the data.

# Review feed-back

## What need to improve:

- There is a small amount of unidentified information that makes the data not 100% clean.
- The data processing code should be written into a function so that it can be used conveniently next time.
- The charts should be more diverse instead of just focusing on columns, so there is a lack of diversity.
- Some data is still unused.
- When making data, it is recommended to use a filter in the range of 100 so that users can easily grasp it.

# Review feed-back

## What can be add for future

- Regarding data cleaning, some other powerful libraries such as structs need to be used to optimize data cleaning.
- Finding many convenient solutions will also help with data cleaning.
- In terms of overall design, the majority of designs are aimed at car auctions and data collection.
- The chart should be designed in many different formats, not focusing too much on the Bar chart.

# BI tools help to make decision-making

BI tools unify multiple data sources, helping to support the overall organization of the business so that managers and employees spend less time tracking information and can focus on generating reports accurate and timely reporting. Armed with up-to-date and accurate information, employees can focus on short- and long-term goals and analyze the impact of their decisions.

Advantage: Saves time, is handy, and information is managed properly.

Disadvantage: BI needs expertise and experience, as well as time to process data.

# BI tools help to make decision-making

## Real company examples:

A corporate sales team can use BI to create a dashboard that shows the company with the highest auction return data and learn their business tactics. Tableau Data Interpretation helps quickly identify possible explanations for outliers and trends in data, allows companies to understand business performance through year-over-year metrics to draw conclusions. results are the car companies with the highest auction value.

# BI tools help extending target audience and competitive within the market.

- **Optimizing Existing Business Model**

BI makes selective data analysis easier, accurate for marketers who have difficulty gathering different sets of information and that are too manual. BI tools help collect diverse customer insights, including behavioral analysis and even existing aspirations. All the data that is procured can then be stitched together to optimize the business model on top of the existing audience.

- **Analyze customer behavior**

BI is a kind of extrapolation solution set that uses AI and ML inputs to understand buyer personality and even customer base that is loyal and rewarding in nature. It is possible to have loyal customers but not as productive as desired.

- **Marketing Optimization**

Modern day marketing is no longer guesswork. Therefore, to stay ahead of the current competition, marketers need to optimize their moves using BI and analytics. Needless to say, this is the age of insights and data-driven decision-making, and in order to make a profit, organizations must abandon instincts and experience in addition to relying solely on facts to make money. complete the work.

# Legal issues related to the exploitation of user data

Below is The GDPR (General Data Protection Regulation) includes six principles that define a new set of rules designed to give citizens control over the personal data stored and processed by businesses.

- Data minimization
- Storage limitations (data retention)
- Lawfulness, fairness, and transparency
- Integrity and confidentiality
- Accuracy
- Purpose limitations

GDPR regulations all focus on what personal data businesses hold about users and how it is handled. It also emphasizes on security and protection of customer data.



# Data minimization

Data minimization ensures that organizations collect only the minimum amount of personal data necessary to fulfill a purpose. This reduces the need for consecutive access to the data source for that additional piece of data.

## Storage limitations (data retention)

In order to comply with data minimization and storage limit guidelines, an enterprise's data controller must ensure that personal data is only stored for a limited period of time.

# Lawfulness, fairness, and transparency

**An individual who is a data subject has certain rights as follows:**

- Right to repair
- Right to be notified
- Access rights
- Right to data portability
- Right to Erase (Right to Forget)
- Right to object
- Right to Restrict Processing
- Rights related to automated (robot) decision-making, including profiling

# Integrity and Confidentiality

To keep data secret, your firm must focus on three key areas.

- Data at rest. Personal data should be secured through the use of secure access restrictions such as roles and passwords.
- Data is being sent. Personal data in transit should be exchanged over secure networks, such as secure file transfer protocol (SFTP), or encrypted.
- Breach of data. Certain sorts of data breaches must be reported by businesses within 72 hours of discovery.

# Accuracy

The data collected needs to be accurate, here are 3 ways to ensure the accuracy of the data.

- Check the data's source.
- Check the veracity of the personal information you collect.
- Use mechanisms to accept and apply challenges to accuracy.

# Purpose Limitation

Collected data needs to have certain usage limits, here are 2 principles when using data:

- Purpose specification - Data should only be acquired for specific, stated, and authorized objectives.
- Compatible use - Data must be processed in a way that is compatible with goals.

# Conclusion

Through Business intelligence, we have learned the knowledge of BI, and how to apply them in practice. We have been learning about data analysis and processing through python language. In addition, we also learned how to design data graphs and analyze the correlation between them. In this Asm 2 report, our team presented the theory of BI and gave some examples. We have clarified BI techniques and analyzed them. We covered some of the tools for BI and how we process the data for them. Finally, we analyzed the data chart and give some conclusion.

Through the business intelligence course, our group think that BI enables organizations to control accurate and effective information from which they may analyze, utilize data, forecast pricing trends and consumer behavior, and identify prospective customers to bring on board. Develop proper business strategies to boost the company's competitiveness.

**Thank you**  
**For Watching**