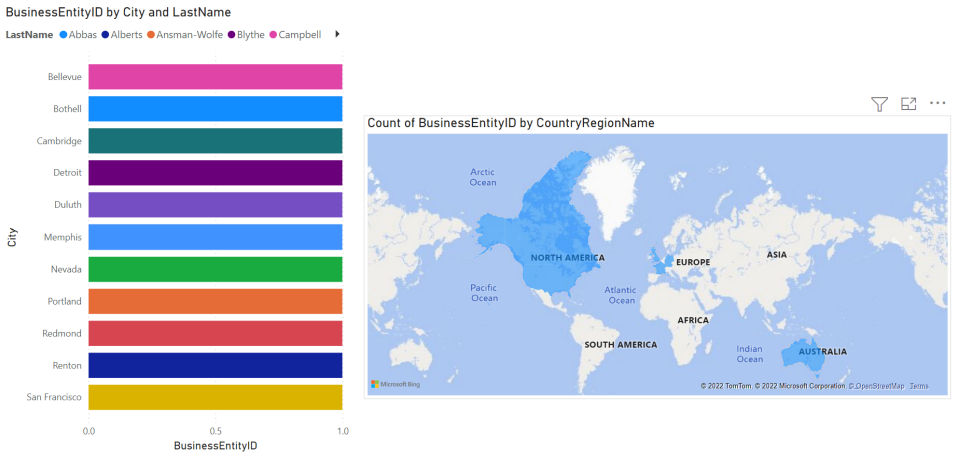
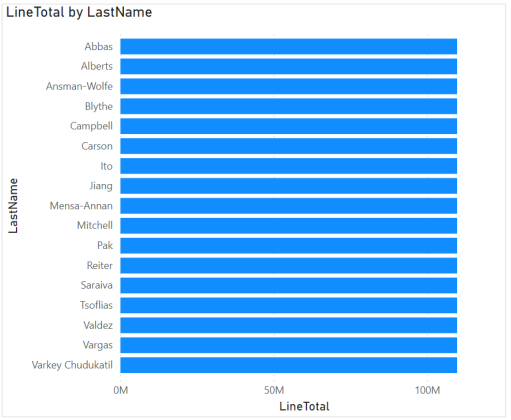
Data Warehouse Lab 3

Task 1

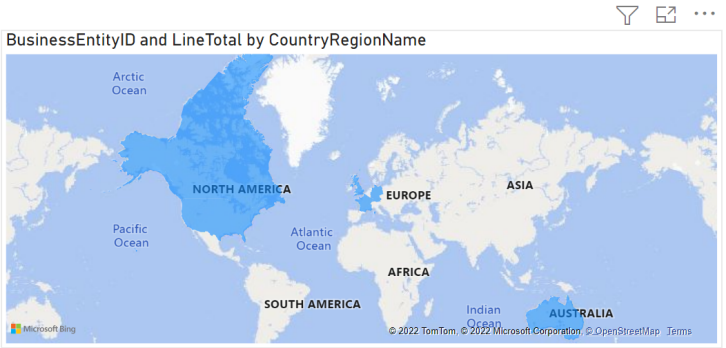
1. A simple report page which shows the sales representatives in a different city and countries *(Report Page Name: Task1-1)*



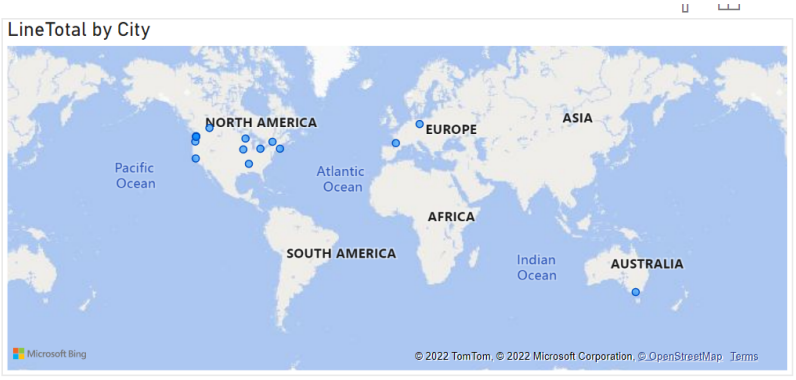
2. b) No because LineTotal is not specific for a Sales Representative, it is more overall for Stores and Sales Representatives. Also because we didn't define the relationship between SalesPerson view and SalesOrderDetail tables.



c) Map on the top right visualizes countries that the sales reps are located (add total sales as tooltip)

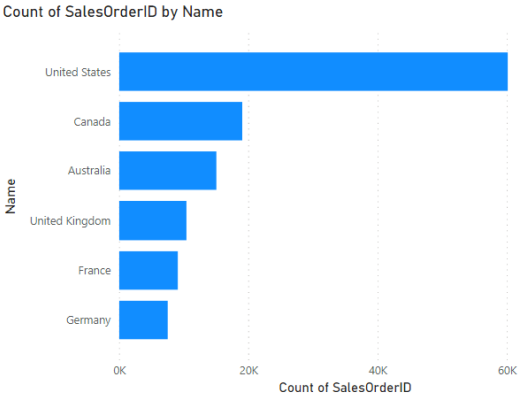


d) Map on the bottom left visualizes total sales (use size) for different cities that the sales representatives are located (use location)



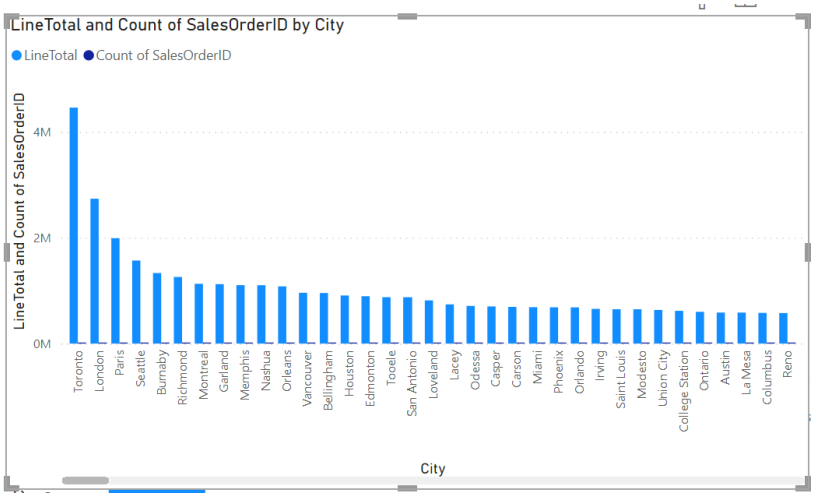
Task 2

2. a)

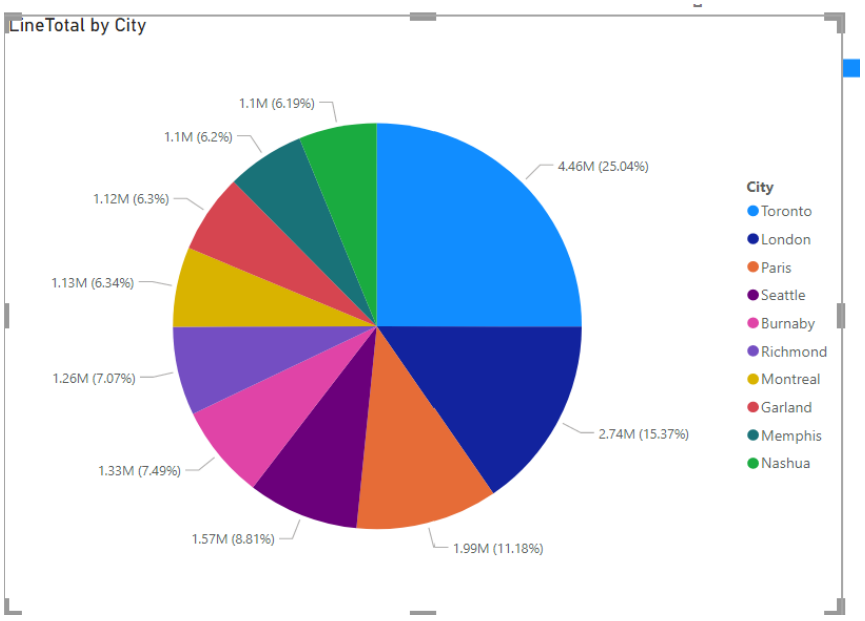


Name	Count of SalesOrderID
Australia	15058
Canada	19064
France	9088
Germany	7528
United Kingdom	10426
United States	60153
Total	**121317**

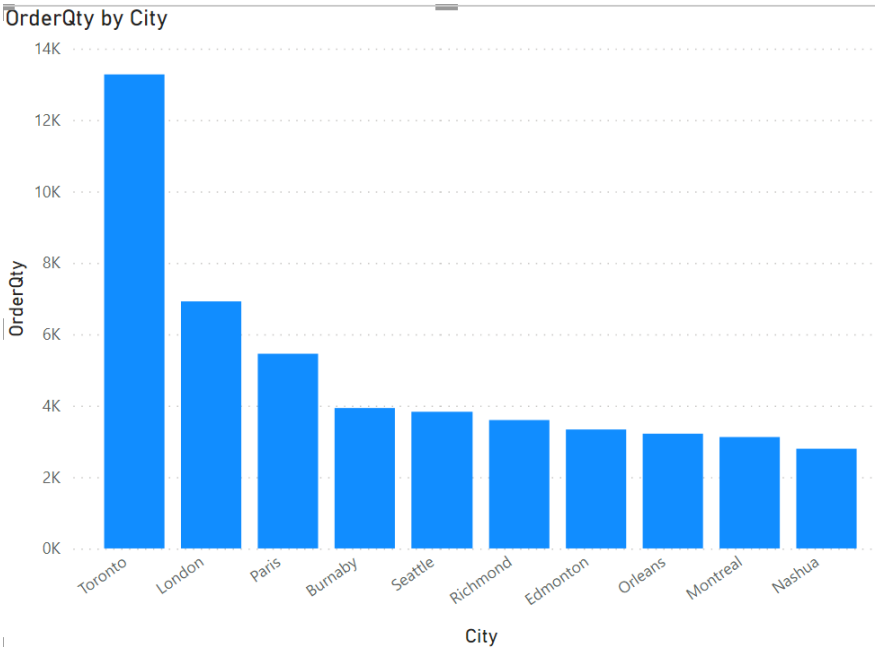
3. a) Present sales values and quantity for different billing locations



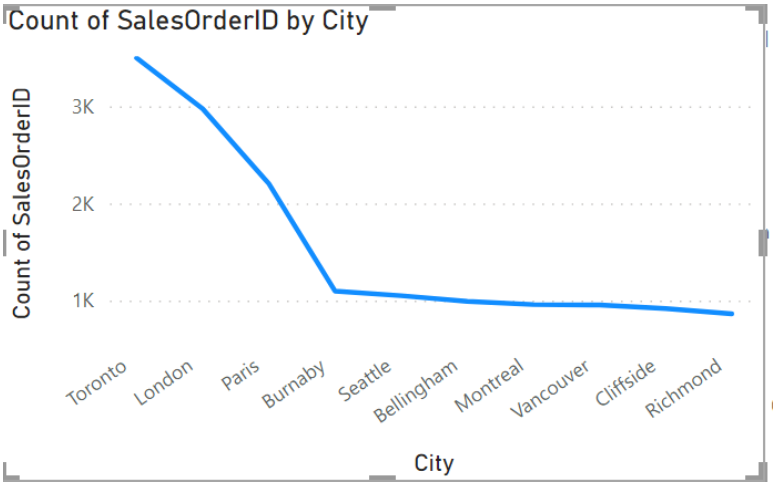
b) Present Top 10 billing cities (sales value wise)



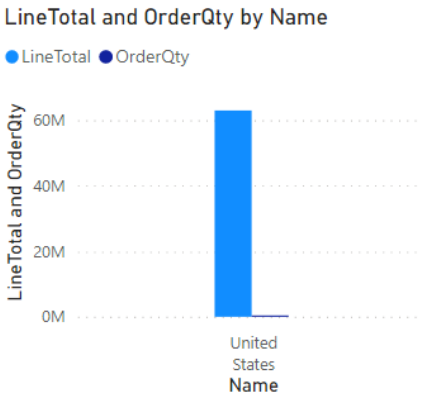
c) Present Top 10 billing cities (sales quantity wise)



d) Present Top 10 billing cities (sales number wise)



e) Present sales values and quantity for top 1 country (sales number wise)



Task 3

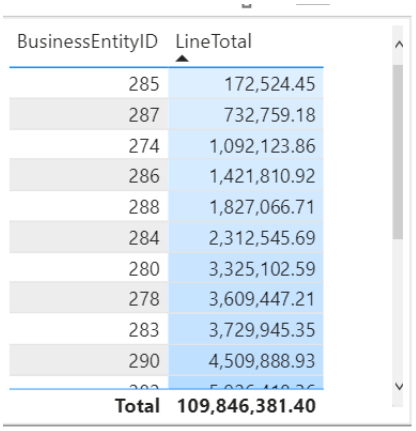
1. a) Present sales values and quantity for different salespersons and different billing location. Consider proper formatting and ordering of data elements.

BusinessEntityID	LineTotal	OrderQty
⊞	**29,358,677.22**	**60398**
⊟ **276**	**10,367,007.43**	**27229**
Tooele	856,562.49	1540
Loveland	816,353.25	2118
Union City	636,226.47	2554
Ontario	600,169.02	1538
Reno	577,089.57	1782
Greeley	572,035.11	1386
Culver City	498,139.72	1378
Newark	456,739.47	862
Gilroy	427,890.84	732
Phoenix	399,011.95	1887
Cerritos	372,146.73	795
Fullerton	349,342.11	1119
Sacramento	338,939.20	539
Lake Elsinore	311,446.43	685
Longmont	309,077.41	1545
Chandler	285,350.20	420
Salt Lake City	273,513.41	613
Modesto	268,209.29	1322
City Of Commerce	261,522.33	263
Fernley	243,014.55	194
Sherman Oaks	182,025.06	258
San Ysidro	140,874.51	86
Mesa	114,294.43	288
Total	**109,846,381.40**	**274914**

b) Identify and present billing countries for top 3 salespersons (sales value wise)

BusinessEntityID	Name
	Australia
	Canada
	France
	Germany
	United Kingdom
	United States
276	United States
277	United States

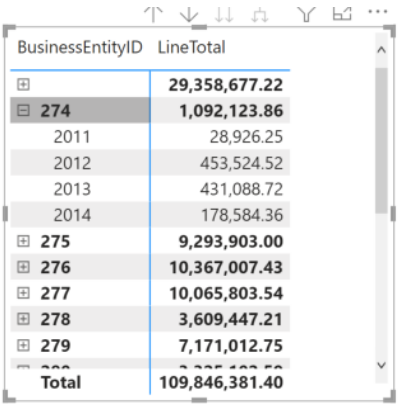
2. a) Use table visualisation with conditional formatting to show salesperson total sales amount.



BusinessEntityID	LineTotal
285	172,524.45
287	732,759.18
274	1,092,123.86
286	1,421,810.92
288	1,827,066.71
284	2,312,545.69
280	3,325,102.59
278	3,609,447.21
283	3,729,945.35
290	4,509,888.93
Total	**109,846,381.40**

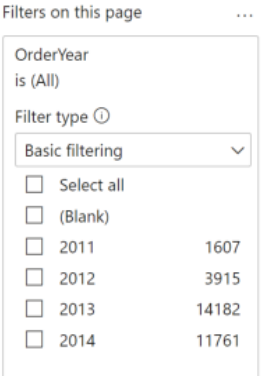
b) Use bar visualisation to show total sales amount by different years – sort the data by sales amount

c) Use matrix visualisation to show total sales by salesperson and year



BusinessEntityID	LineTotal
	29,358,677.22
274	**1,092,123.86**
2011	28,926.25
2012	453,524.52
2013	431,088.72
2014	178,584.36
275	**9,293,903.00**
276	**10,367,007.43**
277	**10,065,803.54**
278	**3,609,447.21**
279	**7,171,012.75**
Total	**109,846,381.40**

d) Use filter – page level – to filter data based on year (order date)

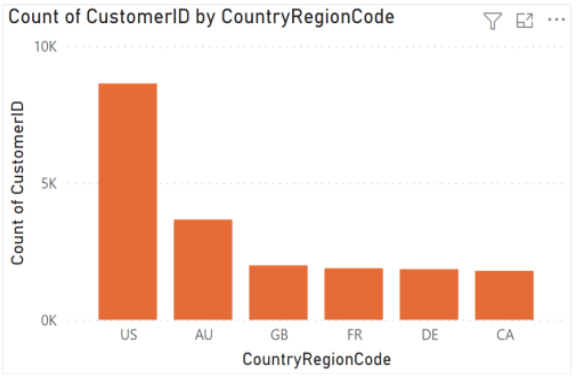


2. Prepare a set of additional 3 visualisations – for each, name the provided information, justify your selection of message (how they support the decisions) and presentation (why this type of visualisation and why such formatting)

1) Column chart visualisation which represents countries with the amount of customers.

It can give some insights like identifying the countries with the most customers, so it can support the decisions like organizing the capacity of work in each region, increasing sales representatives based on the amount of customers, etc.

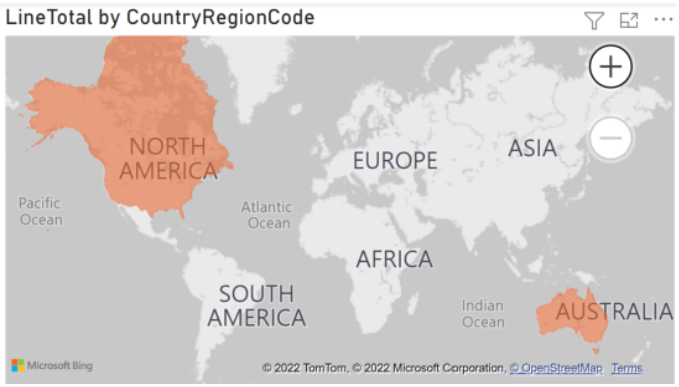
This type of visualization has been selected, because it shows the amount of customers and country regions sorted by the most customers to the less number of customers.



2) Filled map visualization which shows the top 3 countries by sales value.

It can be seen that the top 3 countries by sales value in the last year are the US, Australia, Canada. Based on this information top managers can make decisions such as which country is performing better, and increase sales representatives in those countries. Also to check what can be improved in other countries based on sales value.

This kind of visualization has been selected, because it perfectly shows the location of each country and shows sales value in each country.



3) The third visualization is a matrix table which depicts the top 10 products by sales value and how much quantity was ordered.

Based on this information, there can be some predictions about what kind of products will be sold most next year, and how much quantity will be ordered. It can support decisions such as planning how much pieces of particular product should be made in the future, to analyze the need of the market for customer.

This visualization is perfect to describe such information, because it consists of different columns and shows the total amount, and detailed information if needed.

Name	LineTotal	OrderQty
Mountain-200 Black, 38	1,045,214.64	619
Mountain-200 Silver, 38	912,463.20	508
Mountain-200 Black, 42	909,303.50	513
Mountain-200 Silver, 42	797,855.70	441
Mountain-200 Silver, 46	792,630.15	432
Mountain-200 Black, 46	777,324.13	437
Road-350-W Yellow, 48	677,468.43	586
Touring-1000 Yellow, 60	634,938.87	397
Touring-1000 Blue, 60	634,853.76	393
Road-350-W Yellow, 40	629,707.69	530
Total	**7,811,760.07**	**4856**

Task 4

DOMAIN DATA DICTIONARY

	Location	**Attribute name**	**Attribute type**	**Description**
1	LabData-RatingIPLocationNew	City	Text	A city related to IP Address.
2	LabData-RatingIPLocationNew	Country	Text	A country related to IP Address.
3	LabData-RatingIPLocationNew	CountryCode	Text	A short representation of the country with two letters.
4	LabData-RatingIPLocationNew	Continent	Text	A continent related to IP Address.
5	LabData-RatingIPLocationNew	IP	Text	IP Address where rating has been given.
6	LabData-RatingNew	ReviewID	Numerical	A unique numerical representation of the review
7	LabData-RatingNew	ProductID	Numerical	A unique numerical representation of the product
8	LabData-RatingNew	IP	Text	An IP Address
9	LabData-RatingNew	Date	Date	Date when rating has been created
10	LabData-RatingNew	ratingWebsite	Numerical	rating related to Website
11	LabData-RatingNew	ratingShipping	Numerical	rating related to Shipping
12	LabData-RatingNew	ratingProduct	Numerical	rating related to Product

13	LabData-RatingNew	ratingOverall	Numerical	Overall Rating
14	LabData-RatingNew	gender	Text	Gender of the customer
15	LabData-RatingNew	email	Text	Email of the customer
16	LabData-RatingNew	job	Text	Job title of the customer
17	LabData-RatingNew	postCode	Text	Postal code
18	LabData-RatingNew	source	Text	Source where the survey has been recorded
19	LabData-RatingNew	didPurchase	Text	Shows if the customer purchased the product
20	LabData-RatingNew	didRecommend	Numerical	Shows if the customer recommends
21	LabData-RatingNew	isUsefull	Numerical	Checks if the rating was usefull
22	LabData-RatingNew	userAgent	Text	User Browser and Device

QUALITY ASSESSMENT SHEET

	Location	**Attribute name**	**Attribute type**	**Type of data**	**# unique values**	**# null values**	**Quality assessment**
1	LabData-RatingIPLocationNew	City	text	Nominal	291	0	All clear
2	LabData-RatingIPLocationNew	Country	text	Nominal	1	0	All clear
3	LabData-RatingIPLocationNew	CountryCode	text	Nominal	1	0	All clear
4	LabData-RatingIPLocationNew	Continent	text	Nominal	1	0	All clear
5	LabData-RatingIPLocationNew	IP	text	Interval	501	0	All clear
6	LabData-RatingNew	ReviewID	Int64.Type	Interval	1000	1	Requires null value handling, remove empty field
7	LabData-RatingNew	ProductID	Int64.Type	Interval	26	1	Requires null value handling, remove empty field
8	LabData-RatingNew	IP	text	Interval	134	0	Remove empty field
9	LabData-RatingNew	Date	date	Ordinal	70	1	Requires null value handling, remove empty field

10	LabData-RatingNew	ratingWebsite	number	Interval	0	0	All clear
11	LabData-RatingNew	ratingShipping	number	Interval	0	0	All clear
12	LabData-RatingNew	ratingProduct	number	Interval	1	1	Requires null value handling, remove empty field
13	LabData-RatingNew	ratingOverall	number	Interval	8	0	All clear
14	LabData-RatingNew	gender	text	Nominal	1	0	Remove empty field
15	LabData-RatingNew	email	text	Nominal	1000	0	Remove empty field
16	LabData-RatingNew	job	text	Nominal	24	0	Remove empty fields
17	LabData-RatingNew	postCode	text	Nominal	0	0	Remove all empty fields
18	LabData-RatingNew	source	text	Nominal	1	0	Remove empty field
19	LabData-RatingNew	didPurchase	text	Nominal	0	0	Remove empty fields, change TRUE to 1
20	LabData-RatingNew	didRecommend	Int64.Type	Nominal	0	544	Requires null value handling, remove empty fields
21	LabData-RatingNew	isUsefull	Int64.Type	Nominal	0	801	Requires null value handling, remove empty field
22	LabData-RatingNew	userAgent	text	Nominal	116	0	All clear