ASSIGNMENT : TABLEAU

Data set : World bank CO2 dataset link

1. Between 2005 to 2011, find out which year had the highest CO2(kt) emission ?

Steps to follow to get the answer

- Use **`CO2kt raw data`** table from world_bank_co2 dataset
- Use data interpreter to clean the data and pivot the columns accordingly
- Rename pivot field names as years and rename pivot field values as values
- Add a data source filter for year and select years from 2005 to 2011
- Use fields years and values in the view

2. In the year 2011 what was the CO2 per capita metric ton emission by India

Steps to follow to get the answer-

- Use **`CO2 per capita raw data`** table from world_bank_co2 dataset
- Use data interpreter to clean the data and pivot the columns accordingly
- Rename pivot field names as years and rename pivot field values as values
- Change the data type of values field from string to Number(decimal)
- Drag country name field to column shelf
- Drag values field to rows shelf
- Drag country name to fitler shelf and select India
- Drag years field to filter shelf and select year 2011

ASSIGNMENT : TABLEAU

Sample superstore dataset link

1. Using the orders table from the sample superstore dataset, determine which segment has the highest percentage profit share of total profit

a) Corporate
b) Consumer
c) Home Office

2. Using the orders table from sample superstore dataset, Determine which shipping mode has the highest average profit

a) First class
b) Second class
c) Standard class
d) Same day

3. Using the orders table from the sample superstore dataset, Determine the number of products with discount values greater than equal to 0.75, by using with a bin size of 0.15

a) 300
b) 450
c) 97
d) 3280

ASSIGNMENT : TABLEAU

Sample superstore dataset link

1. Given that, a data value is considered an outlier if it is significantly higher or lower than the other values in the data set. So using the orders table from sample superstore dataset, amongst the customers with negative profit value, the outlier with the highest sales is?

a) **Arthoro prichep**
b) **Sean Miller**
c) **Hunter Lopez**
d) **None of the above**

2. Using the orders table from sample superstore dataset and heat map,

Determine which year and segment are marked as the highest sales and highest profit.

a) **Year-2014 and Segment-Corporate**
b) **Year-2015 and Segment-Home office**
c) **Year-2016 and Segment-Corporate**
d) **Year-2017 and Segment-Consumer**

3. Using the orders table from sample superstore dataset, From the below-given options select product(s) that are in the top 10 in terms of sales value from the furniture category,Using the context filter

a) **Bretford rectangular conference table tops**
b) **SAFCO Arco folding chair**
c) **HON 5400 series task chairs for big and tall**
d) **Aluminum document frame**

4 Find which region has the highest CO2(kt) emission (World bank CO2 dataset link)

a) **North America**
b) **East Asia and Pacific**
c) **Sub Saharan Africa**
d) **Europe and Central Asia**

ASSIGNMENT : TABLEAU

sample superstore dataset link

1. Using the orders table from sample superstore dataset, Find the maximum number of days it took to ship an order from the order date using calculated field and number function (Use calculated fields to know the number of delivery days using ship date & order date)
   a) 2
   b) 5
   c) 6
   d) 7

2. Determine which product subcategory has the highest average profit and the second-highest average cost

Note: Cost= Sales-Profit

a) Machines
b) Tables
c) Bookcases
d) Copiers

3. Find which region has the highest CO2(kt) emission

World bank CO2 dataset link

Steps-

1. Use Co2(kt) Raw data table from the World_Bank_CO2 dataset
2. Clean the data using data interpreter
3. Pivot the years column and rename pivot field names as years and rename pivot field values as CO2(kt)
4. Perform inner join between CO2(kt) raw data table and Metadata-countries table on Country code
5. Add data source filter on Region column and select all regions except null
6. Build a visualisation using fields region and CO2(kt) to get the answer

Note:Do not use CO2(kt) raw data count field

a) North America
b) East Asia and Pacific
c) Sub Saharan Africa
d) Europe and Central Asia

4. Using the order table from sample superstore dataset,

Select correct options that follow steps to find product(s) that have the lowest average profit

Note: Profit can be both positive and negative. A profit less than 0 indicates a loss.

Options:

Step 1) Drag product name field to column shelf -> drag profit field to rows shelf and select average aggregation-> drag profit field to filter shelf ->select average aggregation and click next-> filter range of values that are less than equal to 0-> sort ascending using the profit field

Step 2) Drag product name field to column shelf-> drag profit field to rows shelf and select average aggregation-> sort in ascending order by profit field

Step 3) Drag product name field to row shelf -> drag profit field to column shelf and select average aggregation -> sort in ascending order by profit field

a) Step 2 >> Step 3 >> Step 1
b) Step 1 >> Step 2 >> Step 3
c) Step 3 >> Step 1 >> Step 2

5. Using the order table from sample superstore dataset,

Select the correct ordering of steps to let the users compare sales of top N profitable product subcategories

Steps:

1. Drag the subcategory field to the column shelf and drag the sales field to rows shelf
2. Create a parameter name it top 5 with data type=integer, Display Format = 5, Current Value =5, allowable value=range, min=1, max=5 and step size=1 and click ok->right click on the created parameter->select show parameter
3. Drag the subcategory field to the filter shelf and select use all, then go to the top tab and select by field next instead of integer select the top 5 parameter and by field profit with sum aggregation -> click ok.

a) 1, 2, 3
b) 3, 1, 2
c) 3, 2, 1

6. What is the difference between a dual-axis chart and a combined-axis chart?

a) Dual axis and combined axis are different terms but have the same meaning.
b) Dual axis chart creates two independent axes while a combined axis chart merges two or more measures into a single axis.
c) Combined axis chart creates two independent axes while a dual axis chart merges two or more measures into a single axis.
d) Dual axis chart becomes a combined axis chart once two or more measures are combined into a single axis.

7. _____ is a free platform offered by tableau to explore, create and publicly share data visualisations online.

a) Tableau Desktop
b) Tableau Prep
c) Tableau Public
d) Tableau server

8. We can connect to various file types like Text files, Excel files, etc in Tableau public using different available Connectors.

a) True
b) False

ASSIGNMENT : TABLEAU

1. Using the orders table from sample superstore dataset, Select the correct ordering of steps to find out state-wise percentage contribution to total sales using Fixed LOD

Steps:

1. Create a fixed LOD calculation { FIXED :SUM([Sales]) } that will get the total sales and name it as Total Sales
2. Create a calculated field that computes the percentage sales contribution SUM([Sales])/SUM([Total Sales]) name it % sales
3. Drag state field to rows shelf
4. Drag %sales field to text marks card shelf
5. Click on %sales field and click on format -> in the default tab under number select percentage up to 2 decimal place

(There can be one or more options to choose from the below options)

a) 1,2,3,4,5
b) 3,1,2,4,5
c) 5,4,3,2,1
d) All of the above

2. Using orders table from sample superstore dataset Create a visualisation to determine whether older customers tend to contribute more to sales or not, and then select the correct options

Steps:

1. Create a calculated field name it customer acquisition date-> Enter the formula { FIXED [Customer ID]:MIN([Order Date])}-> click ok
2. Drag the order date field to the column shelf
3. Drag the sales field to the rows shelf
4. Drag the customer acquisition date field to colour marks card shelf->change the mark type to bar
5. Drag the sales field to label marks card shelf-> add a quick table calculator percent of total and compute using table down

Options :

a) The correct ordering of steps is 1 -> 2 -> 3 -> 4 -> 5
b) The correct ordering of steps is 2 -> 3 -> 4 -> 5 -> 1
c) In 2017, customer with acquisition date of 2014 made highest contribution to sales
d) We can conclude that older customers contribute more to sales

3. Using orders table from sample superstore dataset, Create a plot that compares the average sales of each subcategory to the average sales of the respective product category, and select the correct options

Steps:

1. Drag the category and subcategory fields to the rows shelf
2. Create a calculated field name it average sales by category and enter the formula { EXCLUDE [Sub-Category]:AVG([Sales])}
3. Click on show me and select the text table chart
4. Drag measure names field to filter shelf and select only fields sales and average sales by category.
5. Drag measure names filed to the columns shelf.
6. Drag measure values field to text marks card shelf
7. Select average aggregation for the sales field under the measure values area.

a) The average sales for the furnishing subcategory were lower than the average sales for the furniture category.
b) The average sales for the binders subcategory were lower than the average sales of office supplies category.
c) The average sales for the machines subcategory were lower than the average sales of technology category.
d) All the given options.

4. Select the correctly formatted Fixed LOD calculated field

a) FIXED Sub-category :SUM([Sales])
b) {FIXED [Sub-category] : SUM([Sales]) }
c) {FIXED [Sub-category] : [Sales] }
d) All the given options

5. _____ level of detail expressions compute values using the specified dimensions in addition to whatever dimensions are in the view

a) Fixed
b) Include
c) Exclude

ASSIGNMENT - TABLEAU

1. Find the total number of gold medals won by the country India in the year 1928

Data set : Modified summer Olympic dataset link

Steps:

1. Use all medalists table from Modified_Summer_Olympic_medallists_1896-2008 dataset
2. Perform inner join operation between All medalists table and team events fixed all years total table on column NOC
3. Convert the Edition field to dimension
4. Use the country field in the view
5. Using country field perform filter for country=India
6. Using medal field perform filter for medal=gold
7. Using edition field perform filter for year=1928
8. Use count of medals field on text marks card shelf

a) 1
b) 15
c) 5
d) 12

2. Using orders table from sample superstore dataset, Choose the appropriate options, that creates a chart showing which subcategories are losing money in each region

1. Create a calculated field name it profitable or not and use formula IIF(SUM([Profit])<0,'Not profitable','Profitable')->Drag region field to rows shelf->Drag subcategory field to rows shelf->Drag profit field to column shelf->Drag profitable or not to filter shelf and select not profitable
2. Drag region field to rows shelf->Drag subcategory field to rows shelf->Drag profit field to column shelf
3. Create a calculated field name it profitable or not and use formula IIF(SUM([Profit])<0,'Not profitable','Profitable')->Drag region field to rows shelf->Drag subcategory field to rows shelf->Drag profit field to column shelf->Drag profitable or not to colour marks card shelf

a) 1,2,3
b) 1,2
c) 2,3,1

3. What is the difference between a dual-axis chart and a combined-axis chart?
a) Dual axis and combined axis are different terms but have the same meaning.
b) Dual axis chart creates two independent axes while a combined axis chart merges two or more measures into a single axis.
c) Combined axis chart creates two independent axes while a dual axis chart merges two or more measures into a single axis.
d) Dual axis chart becomes a combined axis chart once two or more measures are combined into a single axis.

4. When creating a relationship between two tables in tableau, we must choose between a left, right, or inner join.
a) True
b) False

ASSIGNMENT - TABLEAU

1. Using the orders table from sample superstore dataset, Select the correct ordering of steps to find out state-wise percentage contribution to total sales using Fixed LOD

Sample superstore dataset link

Steps:

1. Create a fixed LOD calculation { FIXED :SUM([Sales]) } that will get the total sales and name it as Total Sales
2. Create a calculated field that computes the percentage sales contribution SUM([Sales])/SUM([Total Sales]) name it % sales
3. Drag state field to rows shelf
4. Drag %sales field to text marks card shelf
5. Click on %sales field and click on format -> in the default tab under number select percentage up to 2 decimal place

a) 1,2,3,4,5
b) 3,1,2,4,5
c) 5,4,3,2,1
d) All of the above

2. For India find which discipline has won the 2nd highest number of medals
Modified Summer Olympic dataset link

Steps:

1. Use All medalist table from Modified_Summer_Olympic_medallists_1896-2008 dataset
2. Perform inner join between All medalist tableandTeam events fixed all years total table on column NOC
3. Build a visualisation using fields, country and discipline.
4. Using country field filter for country India
5. Use the count of medals field on text marks card shelf

a) Hockey
b) Athletics
c) Shooting
d) Boxing

3. Using orders table from sample superstore dataset
Create a visualisation to determine whether older customers tend to contribute more to sales or not, and then select the correct options

Sample superstore dataset link

Steps:

1. Create a calculated field name it customer acquisition date-> Enter the formula { FIXED [Customer ID]:MIN([Order Date])}-> click ok
2. Drag the order date field to the column shelf
3. Drag the sales field to the rows shelf
4. Drag the customer acquisition date field to color marks card shelf->change the mark type to bar
5. Drag the sales field to label marks card shelf-> add a quick table calculator percent of total and compute using table down

a) The correct ordering of steps is 1 -> 2 -> 3 -> 4 -> 5
b) The correct ordering of steps is 2 -> 3 -> 4 -> 5 -> 1
c) In 2017, customer with acquisition date of 2014 made highest contribution to sales
d) We can conclude that older customers contribute more to sales

4. Using orders table from sample superstore dataset and include LOD
From the below given states, determine which state has the highest average top-customer sales.

Sample superstore dataset link

Steps-

1. Create a calculated field that computes the maximum sales value for each customer and apply average aggregation on top of it, name it average of top customer sales.
2. Drag latitude field to rows shelf and longitude field to column shelf->drag state field to details shelf->select marks type as map->click on edit location->click on country->from field select country->click ok->drag state field to label marks card shelf
3. Drag the average of top customer sales field to the color marks card shelf and also to label marks card shelf

a) Wyoming
b) Oregon
c) California
d) Texas

5. Using orders table from sample superstore dataset, Create a plot that compares the average sales of each subcategory to the average sales of the respective product category, and select the correct options

Sample superstore dataset link

Steps:

1. Drag the category and subcategory fields to the rows shelf
2. Create a calculated field name it average sales by category and enter the formula { EXCLUDE [Sub-Category]:AVG([Sales])}
3. Click on show me and select the text table chart
4. Drag measure names field to filter shelf and select only fields sales and average sales by category.
5. Drag measure names filed to the columns shelf.
6. Drag measure values field to text marks card shelf
7. Select average aggregation for the sales field under the measure values area.

a) The average sales for the furnishing subcategory were lower than the average sales for the furniture category.
b) The average sales for the binders subcategory were lower than the average sales of office supplies category.
c) The average sales for the machines subcategory were lower than the average sales of technology category.
d) All the given options.

6. For India find the total number of medals won and CO2 per capita (metric tons) emission

World bank CO2 dataset link

Modified summer Olympic dataset link

Steps:

1. Use CO2 per capita pivoted table from World_Bank_CO2 dataset
2. Add new data source
3. Select Team events fixed all years total table from Modified_Summer_Olympic_medallists_1896-2008 dataset
4. Edit Blend relation and add a custom blend relation between the two data sources on country name
5. Use the country name field and CO2 per capita (metric tons) field from the primary data source [CO2 per capita pivoted table] in the view
6. Filter for country India using Country name field from primary data source [CO2 per capita pivoted table]
7. Use totals field from secondary datasource [Team events fixed all years total table] in the view

a) 50.2 metric tons and 20 medals
b) 38.99 metric tons and 50 medals
c) 38.99 metric tons and 20 medals
d) None of the above

7. Which country got the highest number of silver medals in hockey
Modified summer Olympic dataset link

Steps

1. Use All medalist table from Modified_Summer_Olympic_medallists_1896-2008 dataset
2. Perform inner join between All medalist table and Team events fixed all years total on NOC column to get country names
3. Build a visualization using country field from Team events fixed all years total table and discipline, medal fields from All medalist table.
4. Perform filter for discipline field=Hockey
5. Use count of medals field from All medalist table on text marks card shelf

a) India
b) USA
c) Pakistan
d) Netherlands

8. Select correct options that follow a logical step to get the total number of unique orders returned for each year
sample superstore dataset link

Options:

1. Perform inner join between orders and return tables on order id -> drag order date to column shelf -> drag order id to rows shelf and convert it to measure and select count distinct as an aggregation type -> select marks type as bar
2. Perform inner join between orders and return tables on order id -> drag order date to column shelf -> drag order id to text marks card shelf and convert it to measure and select count distinct as an aggregation type.
3. Perform inner join between orders and return tables on order id -> drag order date to column shelf -> drag order id to rows shelf and convert it to measure and select count as an aggregation type -> select marks type as bar

a) 1,2
b) 3,1
c) 1,2,3
d) 3,2

9. Using the orders table from the sample superstore dataset, Select Correct options that follow logical steps to find the total number of unique products in each subcategory.

sample superstore dataset link

Options:

1. Drag subcategory field to rows shelf -> Drag product name to text marks card shelf -> convert product name to measure and select count distinct aggregation.
2. Drag subcategory to column shelf -> drag product name to details marks card shelf -> create a parameter with integer data type rest all field remains as it is -> create a parameter action where source sheet=current sheet, target parameter=newly created parameter, source field= product name and aggregation =count distinct-> set value= 1 when the selection is cleared-> In the view edit the title and click on Insert and select Parameters.Parameter1 and click ok-> click on any subcategory header in the view.
3. Drag subcategory field to rows shelf-> drag product name to text marks card shelf-> convert product name to measure and select count aggregation.

a) 1,2
b) 2,3
c) 1,2,3

10. Using the orders table from the sample superstore dataset,
Determine the total combined sales for the states of California, Texas, and Washington, as well as the sales amount excluding these states.

sample superstore dataset link

Steps:

1. Create an empty set for the state field and name it the state set
2. In sheet 1 create a map visualization where all the states are displayed (Select mark type as map)
    - Note: If in the map visualization you see unknown then click on it->select edit location->for country/region click on the dropdown and select from field and select country field and click ok.
3. Create another sheet and create a visualization using the state set and sales field.
4. In sheet 1 create change set values action where the source sheet=sheet 1, target set= state set, run action on=select, running the action will= Assign values to set and clearing the selection will= remove all values from set and click on ok
5. Select the above-mentioned states in sheet 1 and go to sheet2 to get the answer.
    - Note:
        - For multi-selection in Windows OS hold down CTRL and select respective states.
        - For multi-selection in MAC OS hold down Command and select respective states.
        -

a) Including states: 766517 & Excluding states: 1530684
b) Including states: 566517 & Excluding states: 1730684
c) Including states: 659984 & Excluding states: 1637217
d) None of the given options

Case Study : Analyse the Data set given below and answer the below questions!
IPL_Cricket dataset (Shared on Whatsapp group)

1. How many unique umpires were there throughout all seasons?
2. Who has won the most man of the match awards?
3. How many games were played in the Bengaluru venue?
4. What is the toss win % of RCB out of all their Tosses?
5. Which city hosted the maximum number of matches and which city has hosted the Minimum number of matches?
6. How many matches were played in the year 2015?