MODUL 3
GOOGLE DATA STUDIO

THEME DESCRIPTION

Students understand and are able to represent distributed big data processing using Google Data Studio to analyze and visualize bigquery data interactively and online.

WEEKLY LEARNING OUTCOMES (SUB-LESSONS)

CLO-1-Sub-CLO-3:

Good understanding of the application of Big Data technology which is the result of efforts in two different fields, like machine learning and cloud computing, -C2.

Through its implementation steps consisting of:

1. Introduction to Google Data Studio
2. Create a data source and bar chart
3. Filtering chart
4. Create a chart using a custom query
5. Visualizing a Bank Account

PRACTICUM SUPPORTS

1. Windows Operating System
2. (any) Browser Application

PRACTICUM STEPS

1) Introduction to Google Data Studio

▸ Data visualization tools can help you make sense of your BigQuery data and help you analyze the data interactively.

▸ You can use visualization tools to help you identify trends, respond to them, and make predictions using your data.
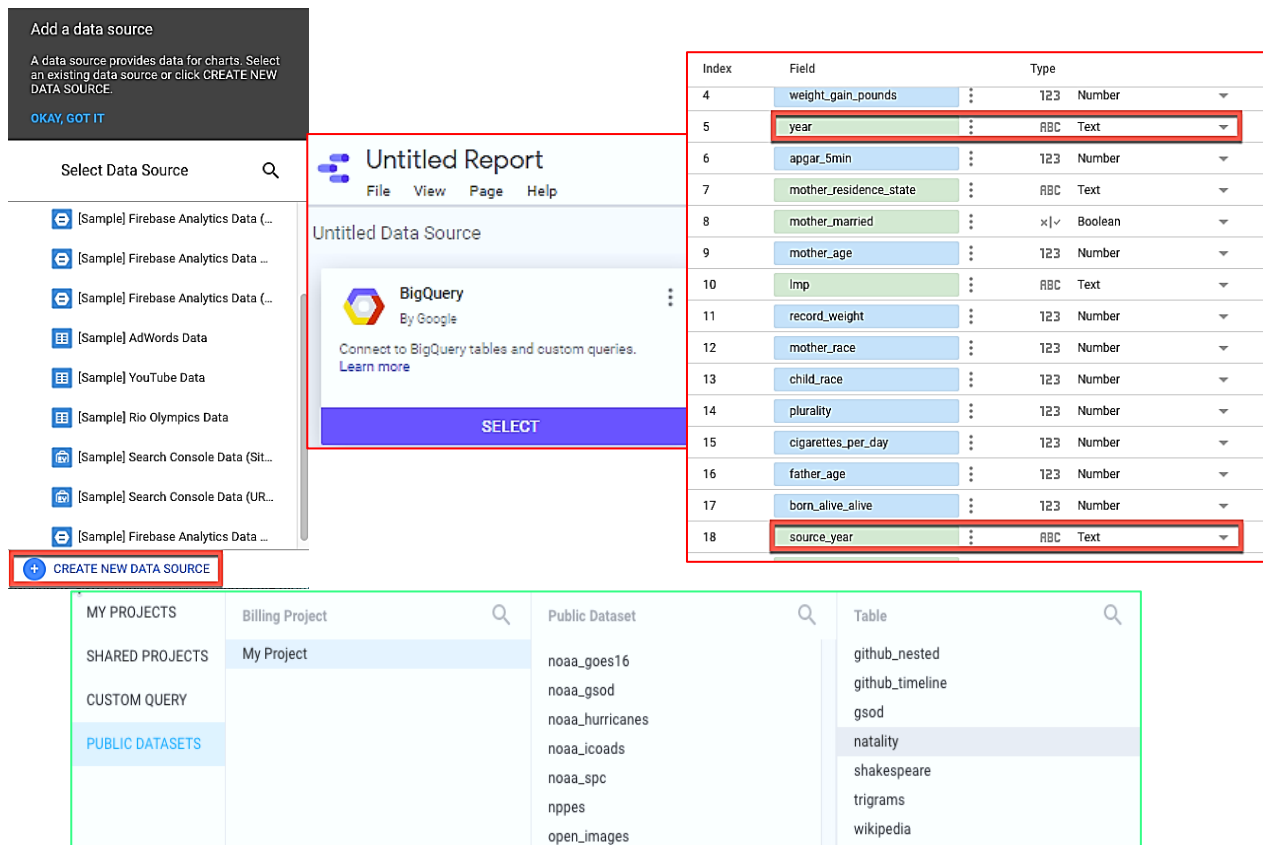

▸ In this tutorial, you use Google Data Studio to visualize data in the BigQuery natality sample table

2) Create a data source and bar chart

▶ The first step in creating a report in Google Data Studio is to create a data source for the report.

▶ A report may contain one or more data sources.

▶ When you create a BigQuery data source, Google Data Studio uses the **BigQuery connector**.

a. Open Google Data Studio: Google Data Studio

b. On the Reports page, in the Start a new report section, click the Blank template. This will creates a new untitled report

c. If prompted, complete the **Marketing Preferences** and the **Account and Privacy** settings and then click **Save**. You may need to click the **Blank** template again after saving your settings.

d. In the **Add a data source** window, click **Create new data source**.

e. In the Google Connectors section, hover over BigQuery and then click Select.

f. For Authorization, click Authorize. This allows Google Data Studio access to your GCP project.

g. In the Request for permission dialog, click Allow to give Google Data Studio the ability to view data in BigQuery. You may not receive this prompt if you previously used Google Data Studio.

h. For My Projects, click Public Datasets.

i. For Billing Project, click the name of the GCP project you created previously.

j. For Public Dataset, click samples for table of Natality.

k. In the upper right corner of the window, click Connect.

▪ Once Google Data Studio connects to the BigQuery data source, the table's fields are displayed.

▪ You can use this page to adjust the field properties or to create new calculated fields.

l. In the natality table, the year and source_year fields are stored as integers.

To use these columns as strings in Google Data Studio, you change the type for these columns to text.

m. To the right of the year and source_year fields, in the Type column, click Number and choose Text from the drop-down list.

n. **Click** Add to report**.**

Create a bar chart using a calculated field

▸ Once you have added the natality data source to the report, the next step is to create a visualization. Begin by creating a bar chart. The bar chart displays the total number of births for each year. To display the births by year, you create a calculated field.

▸ To create a bar chart that displays total births by year:

a. (Optional) At the top of the page, click **Untitled Report** to change the report name. For example, type BigQuery tutorial.

b. After the report editor loads, click **Insert > Bar chart**.

c. Using the handle, place the chart on your report.

d. On the **Data** tab, notice the value for Data Source (natality) and the default values for Dimension and Metric.

e. Because you are charting the number of births by year, the dimension should be set to source_year. If source_year is not the default dimension, click and drag **source_year** from **Available Fields** onto the default dimension to replace it.

f. In the **Metric** section, click **Add metric**.

g. In the metric picker, click **Create field**.

h. To display a count of the number of children born each year (by gender), you create a calculated field. For this tutorial, you count the entries in the is_male field. The value in is_male is true if the child is male and false if the child is female. In the new metric window, for **Name**, type birth_count.

i. For **Formula**, type the following: COUNT(is_male).
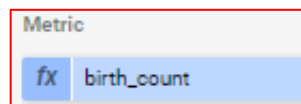
Create a bar chart using a calculated field

Once you have added the data source to the report, the next step is to create a visualization. Begin by creating a bar chart
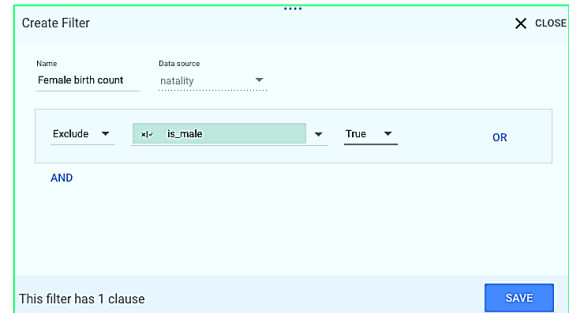
a. Click Apply.

b. Click the report page close the metric picker.

c. After the metric is added, hover over the default metric and click the delete icon on the right hand side.
    The Dimension should now be set to source_year and the Metric should be set to birth_count.

d. Notice the chart is sorted in Descending order by default. The years with the most births are displayed first.

e. To enhance the chart, change the bar labels. In the bar chart properties window, click the Style tab.

f. In the Bar chart section, check Show data labels

g. Sort the data by source year instead of **birth_count**. In the bar chart properties window, on the Data tab, in the Sort section:

h.   Click New field.

i.   In the sort picker, scroll to Chart Fields and select **source_year.**

j.   Click away from the dialog to close the sort picker.

k.   Click Descending and select Ascending.

l.   The display changes to show the number of births by year in ascending order

m. Screenshot your visualization with the display form as Output Figure 3.1 Birth Count by Year

3) **Filtering chart**

Currently, the bar chart displays the total number of male and female children born each year. Add a filter to display only the female children born each year.

a.   In the bar chart properties window, click the **Data** tab.

b.   In the **Filter** section, click **Add a filter**.

c.   In the **Create filter** window:

d.   For **Name**, type Female birth count.

e.   Click **Include** and change it to **Exclude**.

f.   Click **Select a field** and choose is_male.

g.   Click **Select a condition** and choose **True**.

h.   Click **Save**, and screenshot your filter result as shown on Output Figure 3.2 Female children born/year

4) **Create a chart using a custom query**

▸   Creating a chart using the Custom Query option allows you to leverage BigQuery's full query capabilities such as joins, unions, and analytical functions.

▸   When you specify a SQL query as your BigQuery data source, the results of the query are in table format, which becomes the field definition (schema) for your data source.

▸   When you use a custom query as a data source, Google Data Studio uses your SQL as an inner select statement for each generated query to BigQuery.

▸   For more information on custom queries in Google Data Studio, consult the online help.

Add a bar chart using a custom query

▸   To add a bar chart to your report that uses a custom query data source:

a.   From the menu options, choose **Page > New page**.

b.   Click **Insert > Bar chart**.

c.   Using the handle, place the chart on the report.

d.   In the **Bar chart properties** window, on the **Data** tab, notice the **Data Source**

e.   is set to natality. Click natality to open the **Select Data Source** window.

f.   Click **Create new data source**.

g.   In the **Google Connectors** section, hover over **BigQuery** and the click **Select**.

h.   For **My Projects**, click **Custom query**.

i.   For **Project,** select your GCP project.

j. Enter the following standard SQL query in the **Enter custom query** window:

```
SELECT
  source_year,
  sum(case when is_male then 1 else 0 end) as male_births,
  sum(case when is_male then 0 else 1 end) as female_births
FROM
  `bigquery-public-data.samples.natality`
GROUP BY
  source_year
ORDER BY
  source_year
DESC
```

This query uses a CASE statement to evaluate the is_male boolean field to determine how many births were female and how many were male for a given year.

k. Under Query Options, verify Use Legacy SQL is deselected. This allows you to use BigQuery standard SQL syntax.

MY PROJECTS	Project		Enter Custom Query
SHARED PROJECTS	My Project		1 SELECT
			2 source_year,
CUSTOM QUERY			3 sum(case when is_male then 1 else 0 end) as male_births,
			4 sum(case when is_male then 0 else 1 end) as female_births
PUBLIC DATASETS			5 FROM
			6 `bigquery-public-data.samples.natality`
			7 GROUP BY
			8 source_year
			9 ORDER BY
			10 source_year
			11 DESC

Query Options

☐ Use Legacy SQL (?)

Advanced Options ▼

l. At the top of the window, click **Untitled data source**, change the data source name to Male female case query.

m. In the upper right corner of the window, click **Connect**. After Google Data Studio connects to the BigQuery data source, the results of the query are used to determine the table schema.

n. When the schema is displayed, notice the type and aggregation for each field. By default, source_year is formatted as a Number and the Aggregation is set to None. To the right of the source_year field, in the **Type** column, click **Number** and select **Text**. Verify Aggregation is set to None.

| Index	Field		Type		Aggregation
1	Record Count	⋮	123 Number	▼	Auto
2	source_year	⋮	ABC Text	▼	None
3	male_births	⋮	123 Number	▼	None ▼
4	female_births	⋮	123 Number	▼	None ▼

o. Click Add to report. When prompted, click Add to report.

p. Click the back arrow to close the Select Data Source window.

q. Google Data Studio may be unable to correctly determine the appropriate Metric for the chart. Because you are charting the number of male and female births by year, the values for Metric should be

female_births and male_births. In the bar chart properties window, on the Data tab, in the Metric section, click Record Count.

r. In the metric picker, select **female_births**.

s. In the **Metric** section, click **Add metric**.

t. In the metric picker, select **male_births**.

Your chart now displays the number of male and female children born each year using separate bars.

u. The chart should be sorted by female_births in descending order. Change this to source_year as you did in the previous chart. In the Sort section, click female_births.

v. In the sort picker, choose source_year. The chart now shows female and male births each year in descending order.

w. For readability, change the chart styles. In the Bar chart properties, click the Style tab.

x. Notice each bar has a default color based on the order the metrics were added to the chart.

y. Click the second square in the color palette and change it from red to blue.
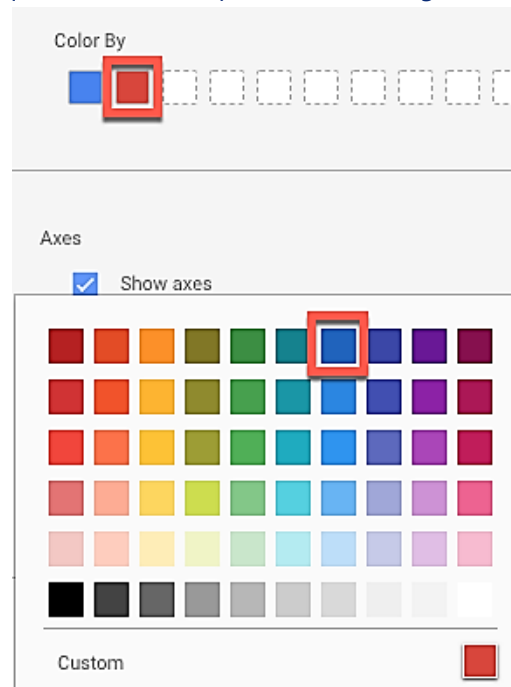
z. Now your visualization clearly distinguishes the color of the bars between female and male babies as shown on Output Figure 3.3 Baby born male & female.

❖ Save all your screenshot(s) result as Figure 5.1 until 5.3 into <mark>IS-429 Lab Week#3 Google DataStudio yourname-NIM.doc/docx</mark>
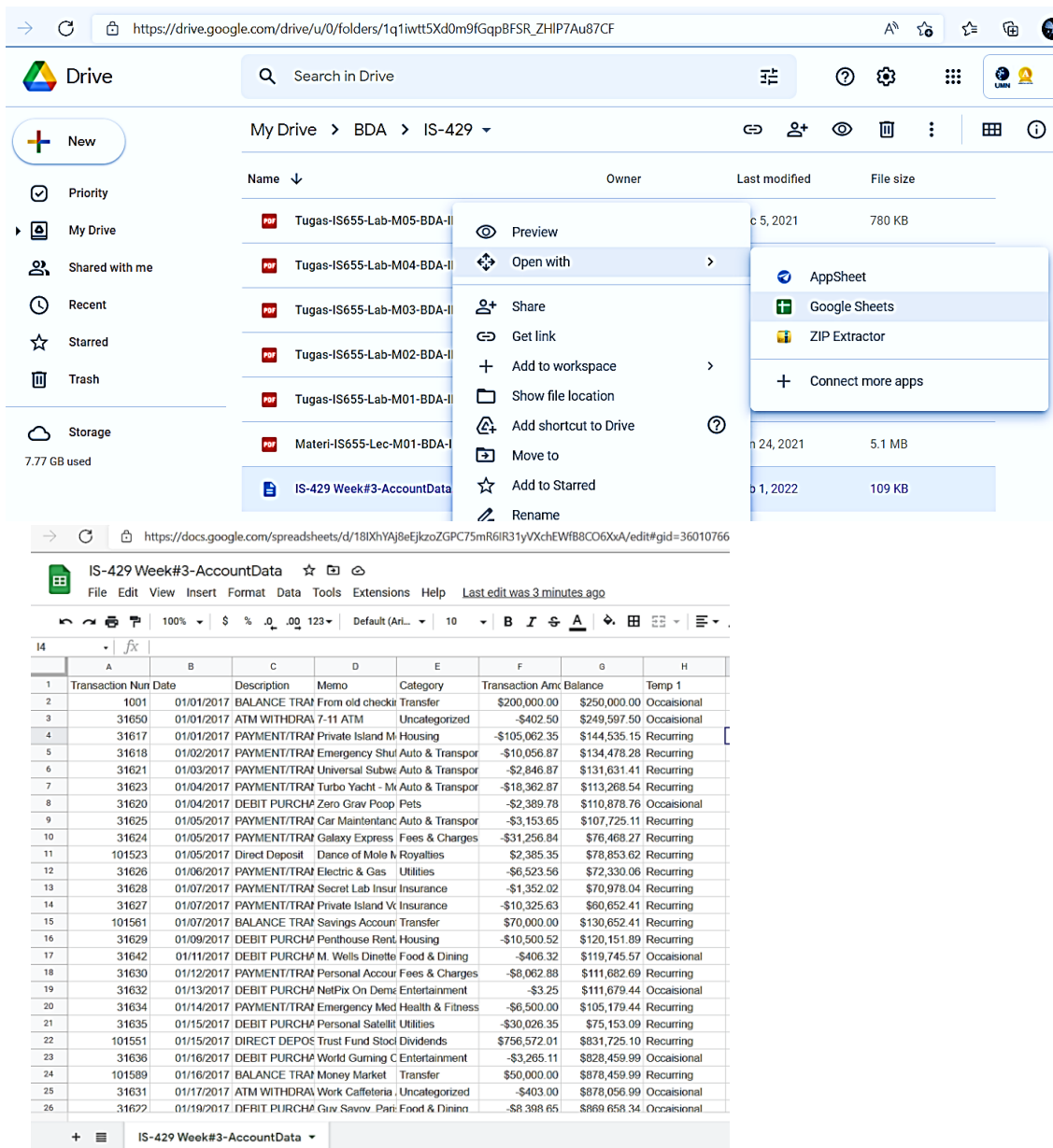
5) Visualizing a Bank Account

This exercise will creates a complete graphical report on checking a bank account which consists of several components such as:

▸ The title of the report is equipped with an image
▸ Date range of report details
▸ Scorecard summary, maximum and minimum balance
▸ Graphical time-series
▸ Transaction details in tabular form

Preparing the Ingredients: Working with Google Sheets

▸ In this case you prepare your Google drive to accommodate IS-429 csv data that will be used in Google Data Studio.

a. Login to your Google drive and upload data file "IS-429 Week#3-AccountData.csv" (*download from e-Learning IS-429 Practicum Week#3*)

b. Save it your csv data into Google sheet.

c. Open Google Data Studio: Google Data Studio

d. On the Reports page, in the Start a new report section, click the Blank template. This will create a new untitled report

e. Upload your Google sheet data onto Google Studio
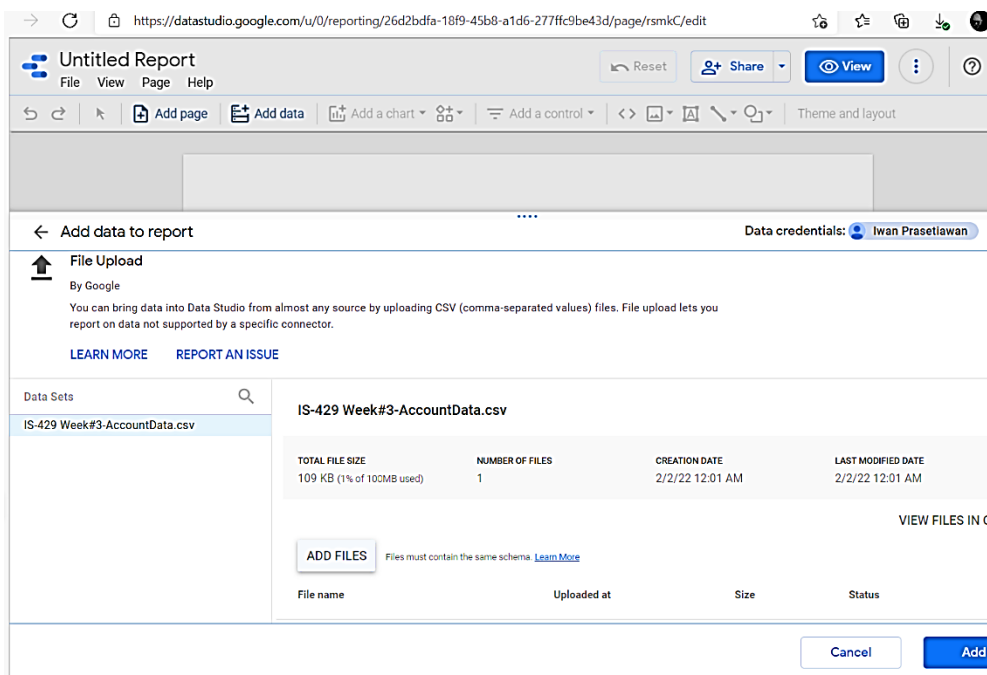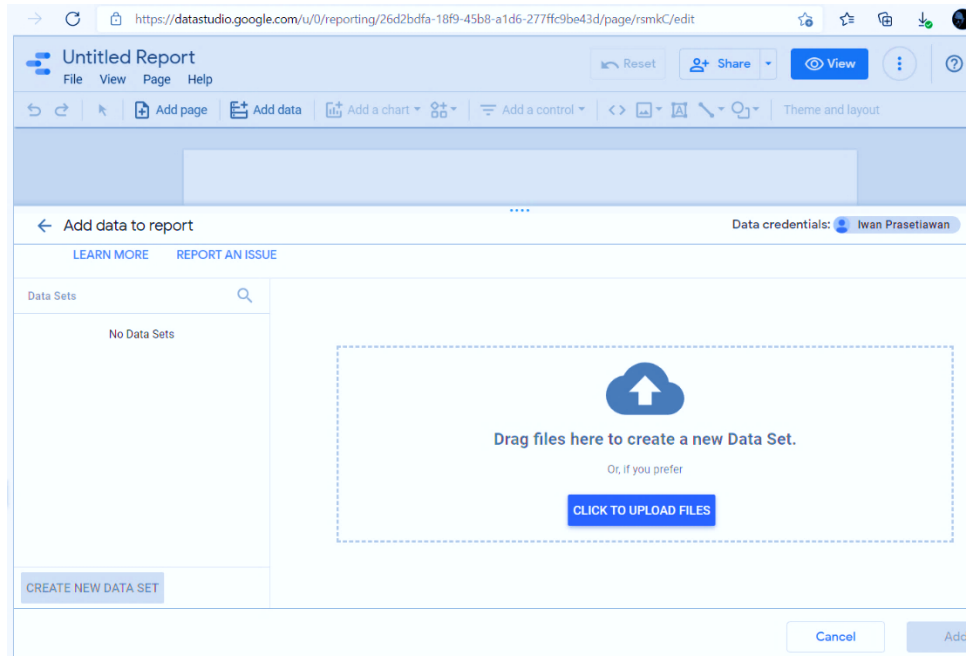
Set the field type and aggregation method for the Transaction Amount and Balance fields.

‣ In the case of Balance, using the Sum aggregation method does not make sense. If you have $50 in an account on Monday and $60 in that account on Tuesday, adding these together doesn't mean that you magically have $110!

a. Setup your data input
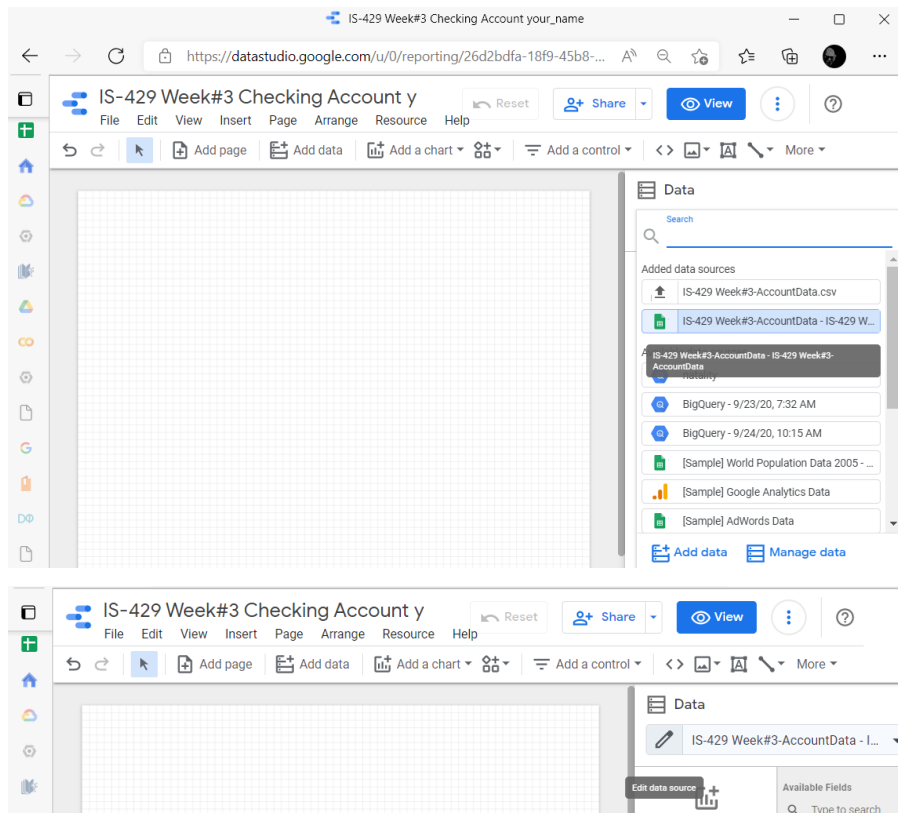
b. In this case, you are better off using Average as the aggregation method, which yields an average balance of $55 over the two days. Thus, for the Balance field, we will set the aggregation method to Average.

c. Save your report page onto "IS-429 Lab Week#3B Bank Account yourName-NIM"

▸ Let's start off with a basic table. Tables are always a good place to begin when creating your reports, even if they don't make it into the final version.

▸ Tables are also a good way to check that all of the data is coming through as you would expect.
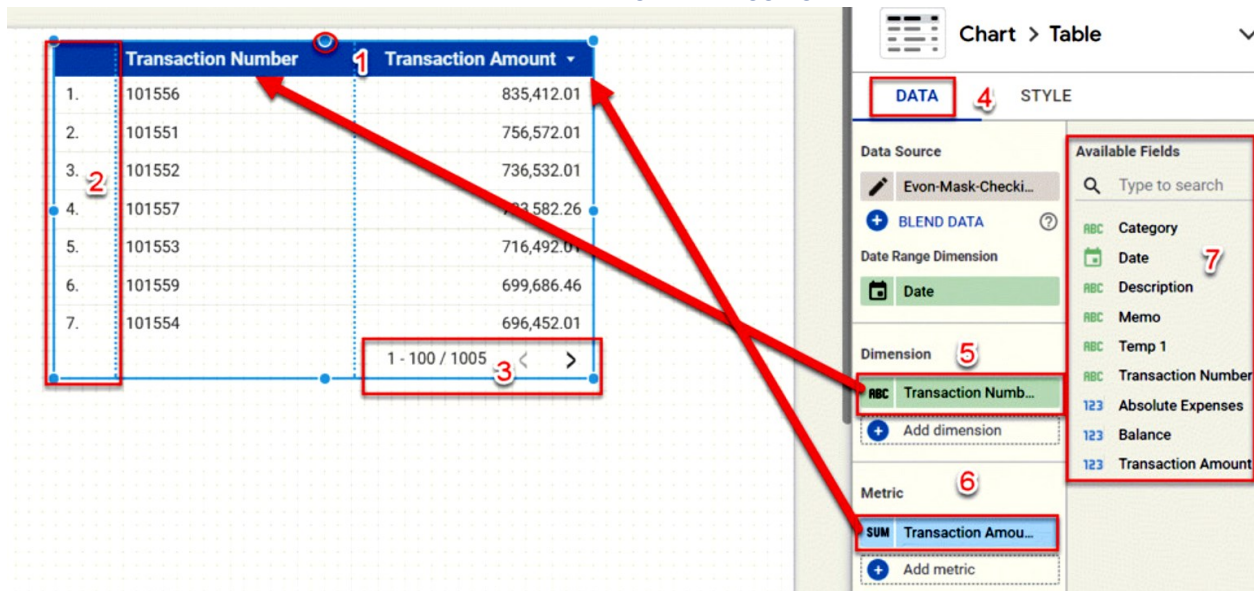
    a. Setting the size of the page canvas.
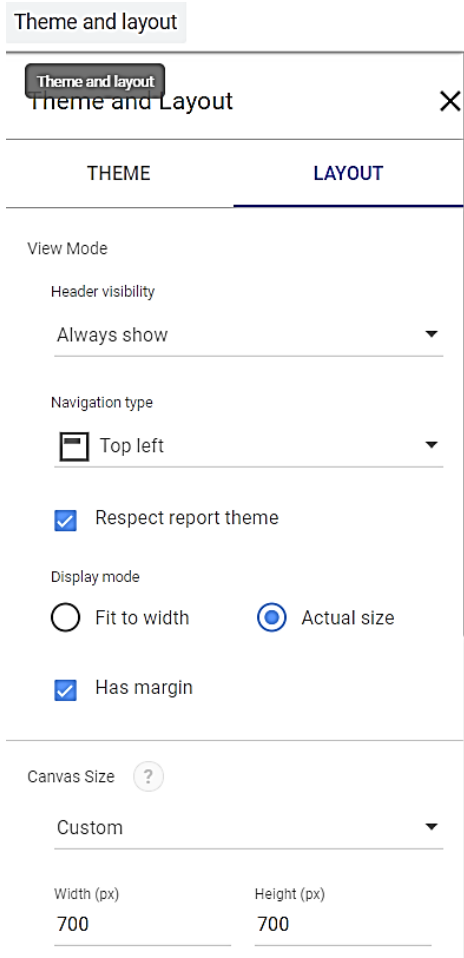
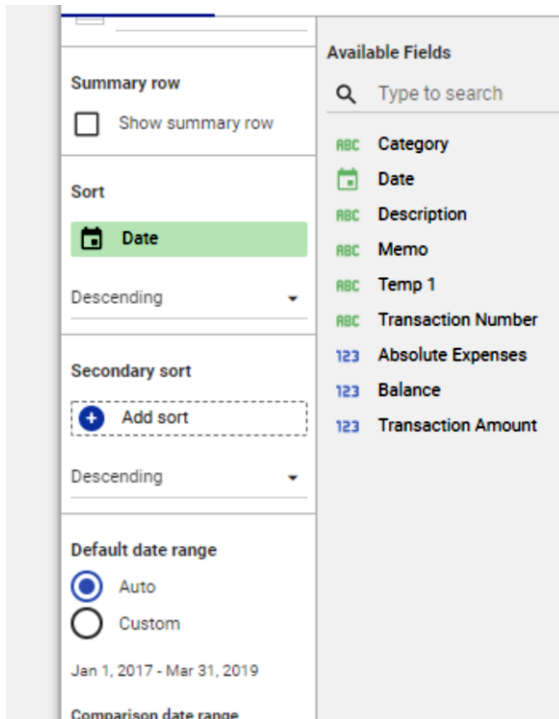   b. Click the Add A Chart drop-down menu, as shown in below Figur and we will use a simple table
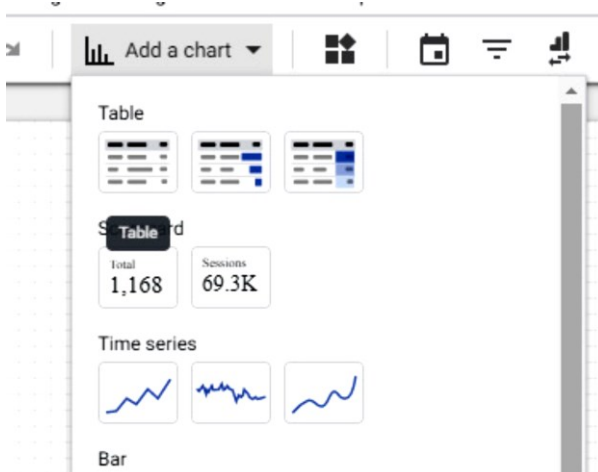
    c. **Table selection**: Once the table is placed on the canvas, it will be selected. You can tell this by the outline and the little dots, or handles, that appear on the outside of the table.

When selected, the table can be moved by clicking and dragging with your mouse, or you can use the
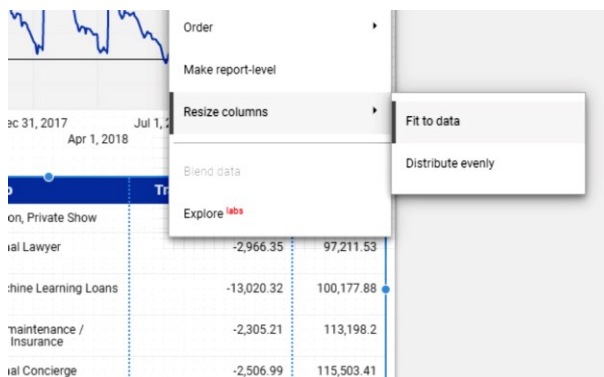

d. Change the sort order so that the items in the table appear with the most recent date at the top.

e. Now we have our table on the page, but it needs some help to make it look better. First, we need to adjust the columns on the table; then we will change the style settings.

f.   We have to make a couple of style adjustments to complete our table. As is often the case, we want to remove the left Number column and the pagination control at the bottom of the table.

  ▸ These controls are located on the Style tab. Unless Google changes the default style options for tables in the future, you should get used to removing the numbering and pagination style settings when you first create a table.

  ▸ To adjust the style, select the table with the mouse and then select the Style tab in the control panel. Scroll down the tab and deselect the boxes for Row Numbers under Table Body and Show Pagination under Table Footer, as shown in below Figure:
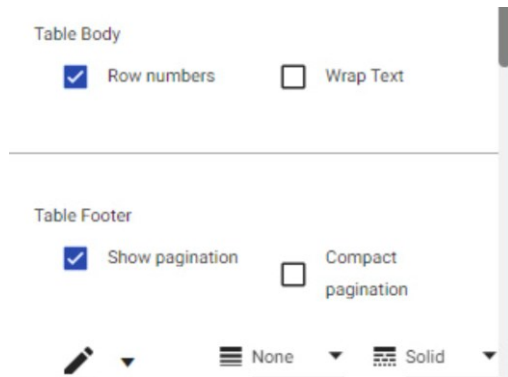
  ▸ Your table should look much neater now. If you wish to, move the table to the bottom of the canvas by clicking and dragging it with the mouse or by using the arrow keys.

  ▸ Click the View button to see how it looks after you move it. It should look similar to our example report

❖ Congratulations on adding your first table!

Keep Score with Scorecards

▸ Now we are going to jump to the top of the report and look at the big numbers in the boxes. This is a very simple but useful component called a scorecard.

▸ Scorecards help us see the big picture. They are primarily used to summarize data over the entire time frame of the report.

▸ Now we are going to the top of the report to create scorecards

a.   Switch to Edit mode by clicking the Edit button in the header.

b.   Select the Scorecard chart from the Add A Chart drop-down menu, as shown in below Figure. Click the canvas to place it.
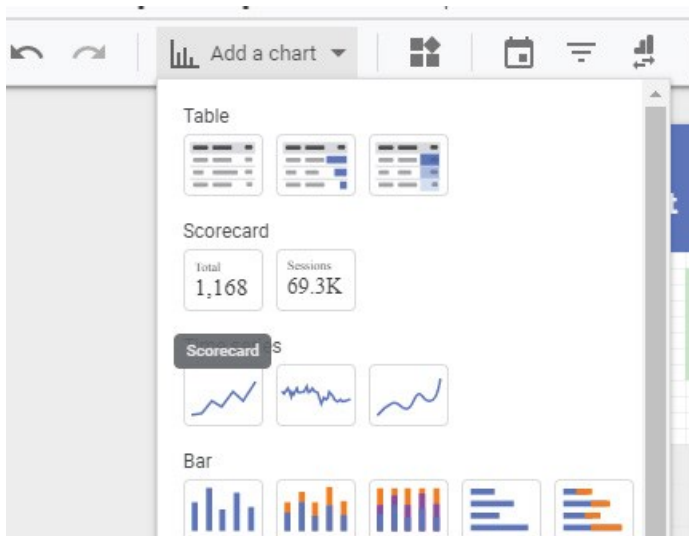
c. On the Data tab, change the metric by dragging the Balance field to replace the Transaction Amount field:
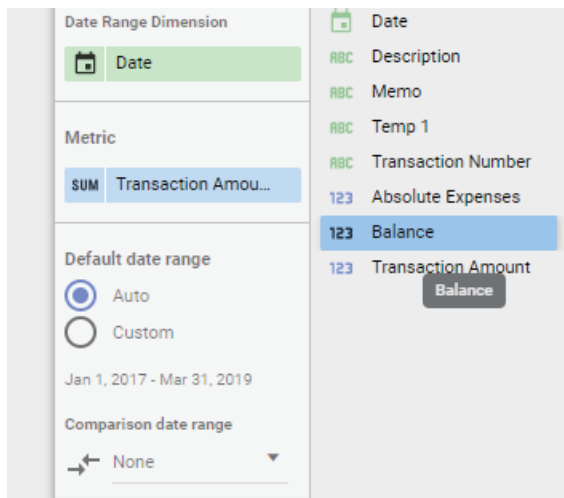
d. Click the pencil icon to edit the field in the chart. A new chart field attribute box will appear. Change the label to Average Balance.
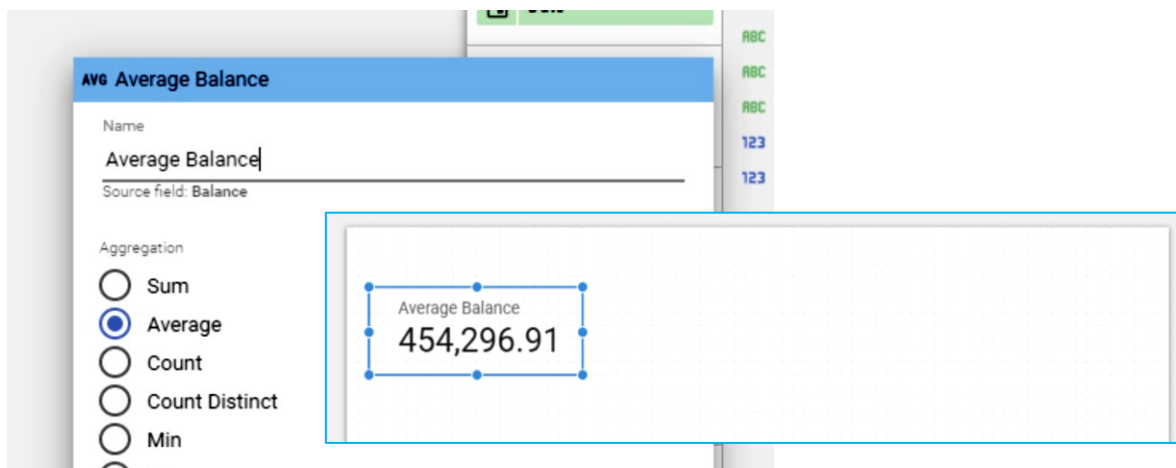
e. Select the Average Balance scorecard. . Copy it to the clipboard using Edit ⇨ Copy, or by pressing Ctrl+C.

f. Paste the scorecard using Edit ⇨ Paste, or by pressing Ctrl+V.
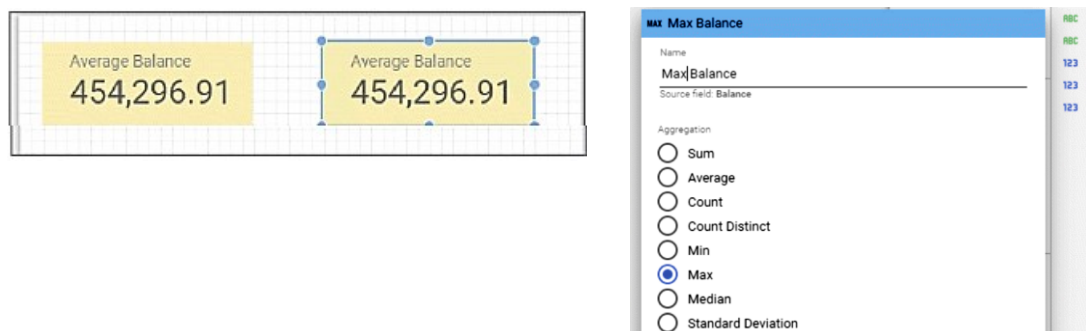
‣ Now click the Data tab. Hover your mouse pointer over the AVG label in the Average Balance field in the Metric section and click the pencil icon.

‣ Change the name to Max Balance, and change the Aggregation setting to Max.

g. Create also for Min. balance scorecard:

❖ Congratulations on completing the scorecard setup.

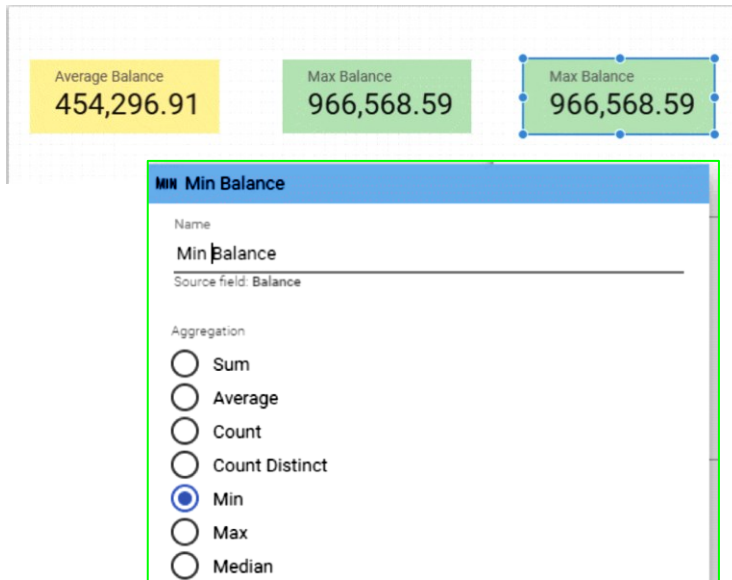| Date ▾	Description	Memo	Transaction Amo…	Balance
Mar 26, 2019	DEBIT PURCHASE	Hamilton, Private Sh…	-25,006.35	72,205.18
Mar 26, 2019	PAYMENT/TRANSF…	Personal Lawyer	-2,966.35	97,211.53
Mar 25, 2019	PAYMENT/TRANSF…	AI Machine Learning…	-13,020.32	100,177.88
Mar 25, 2019	PAYMENT/TRANSF…	Clone maintenance …	-2,305.21	113,198.2
Mar 24, 2019	PAYMENT/TRANSF…	Personal Concierge	-2,506.99	115,503.41
Mar 24, 2019	PAYMENT/TRANSF…	Personal Chef/Nutrit…	-8,659.48	118,010.4
Mar 23, 2019	PAYMENT/TRANSF…	Personal Stylist (par…	-5,361.32	126,669.88

❖ For the final chart, we will examine how to visualize changes over time.

Building the Time-Series Chart

▸ Time-series charts give the viewer a sensory way to grasp trends and patterns in a truly unique way.

▸ We can look back at the past and recognize the importance of things that were not clear to us at any particular moment.

▸ Time-series charts help us understand, in a very visceral way, the story that the data has to tell us. This is the great power of data visualization.

▸ Now let's take a look at your checking account balance chart over 27 months, as shown in below Figure:
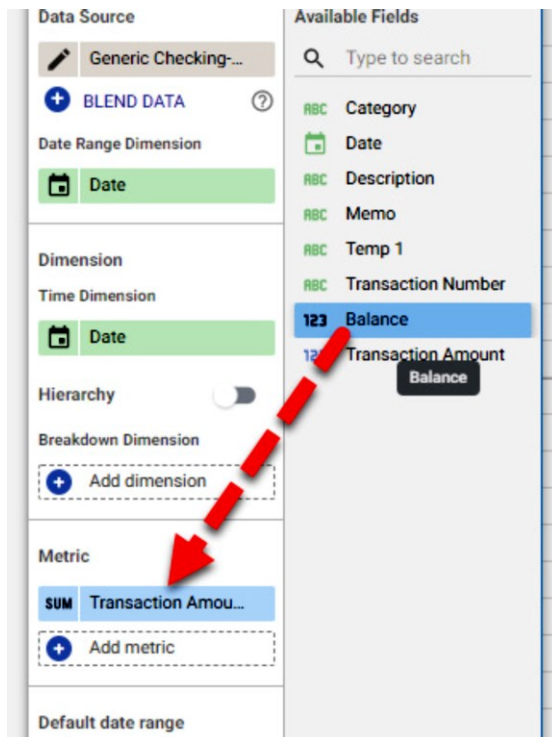

a. In View mode, switch to Edit mode by clicking the Edit button.

b. Select the simple time-series chart from the Add A Chart drop-down. In this case, we will select Smoothed Time Series Chart.
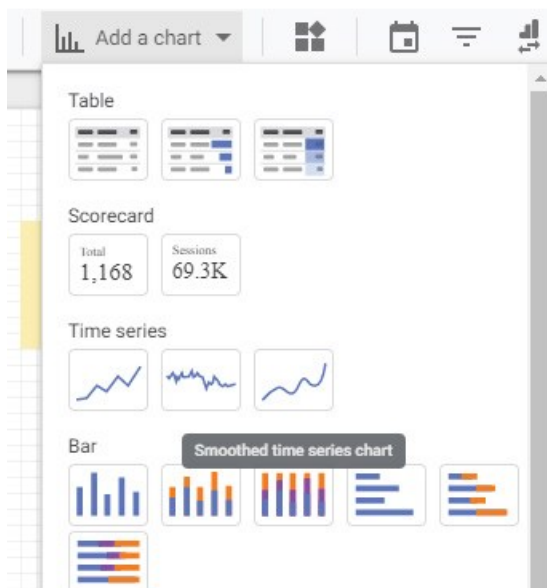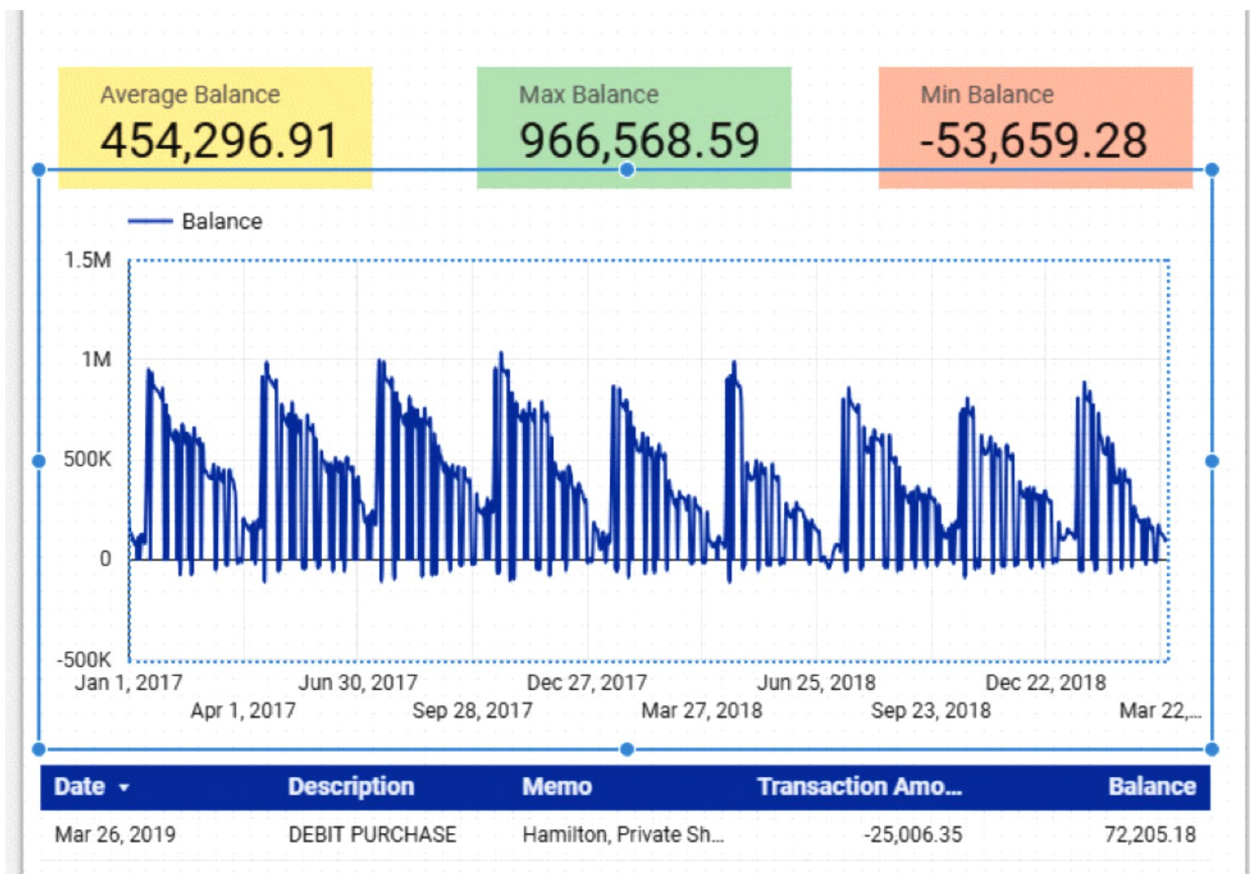
c. On the Data tab, set the metric to Balance by dragging Balance from the Available Fields list:

d. Now position and stretch the chart to fill the middle space on the canvas so that it looks like below Figure:

In the next steps, we will finish our first report by adding and adapting some graphic and interactive elements

Finishing Touches and Sharing Your Report

▸ To give your report a little more style, we will build a header with some graphics.

▸ Regular graphics in Data Studio, shapes, and images work very similar to presentation software like Google Slides or Microsoft PowerPoint.

    a. Adding a header bar: Click the rectangle icon on the toolbar.

        ▸   Your cursor will turn into crosshairs and allow you to click and drag a rectangle shape.

      ▸ The default color here is blue, which is fine for our purposes. Use the arrow keys and mouse to position and size the bar as the header background.
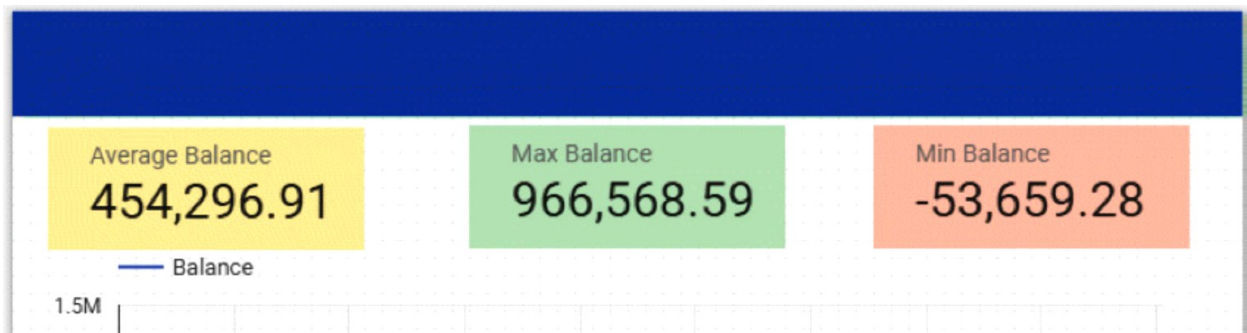
    b. Adding a text box with the report title: Click the text box icon on the toolbar.

      ▸ As with the rectangle, the pointer will turn into crosshairs, so you click and drag with the mouse to create a new text box.

      ▸ On the right side, you will see a Text Properties panel that lets you set the text styles
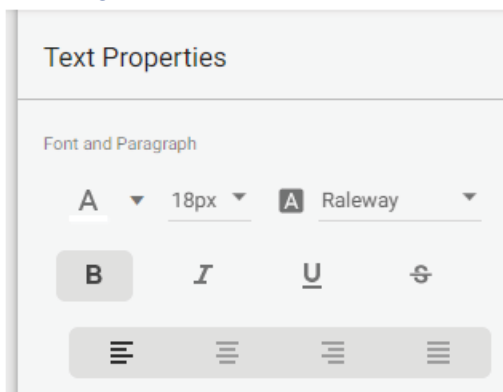
    c. Continue the final completion of the Check Bank Account statement and according to the output figure 3.4 Bank Account Visualization and Save your report.

    d. Screenshotted your final report into "IS-429 Lab Week#3B Bank Account yourName-NIM.jpg"

    e. Data Studio is the ease with which you can share a report, and the variety of sharing and security options available.

      ▸ In this case, we want to share our report so that anyone can view and use the chart. However, only the owner should be able to edit or modify it.

      ▸ Click the Share drop-down above the toolbar next to the View button. Select Get Report Link from the menu.
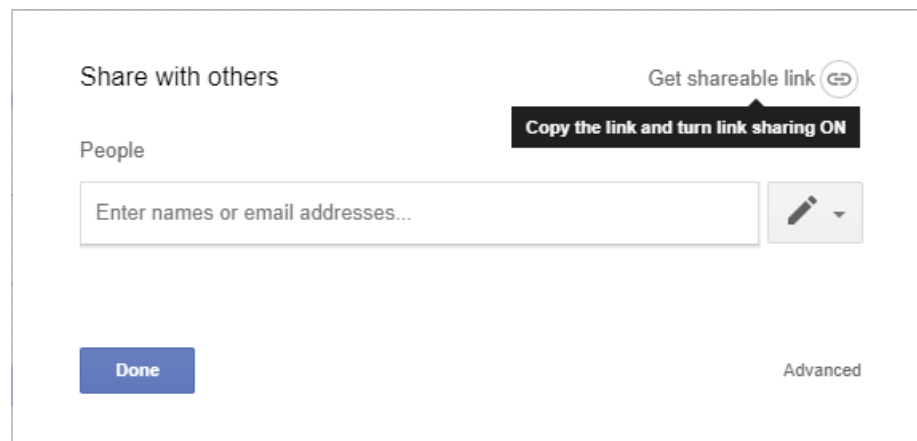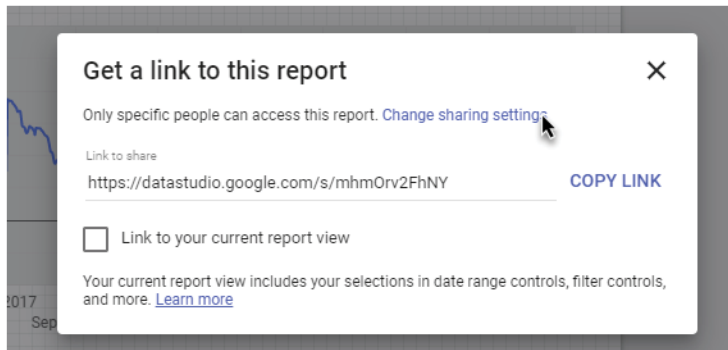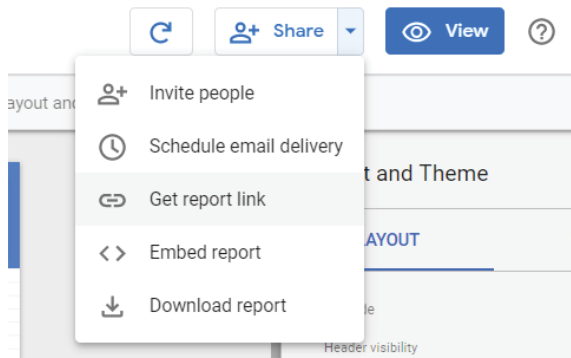
f. "**Get report link**" of your report and save it into "<mark>IS-429 Lab Week#3C Bank Checking report.txt</mark>"

✪Finally, zipped your all of your results (doc/docx, jpg and txt) into "<mark>IS-429 Lab Week#3 Google Data Studio NIM-yourname.zip</mark>" and submit to e-Learning IS-429 Week#3 practicum.
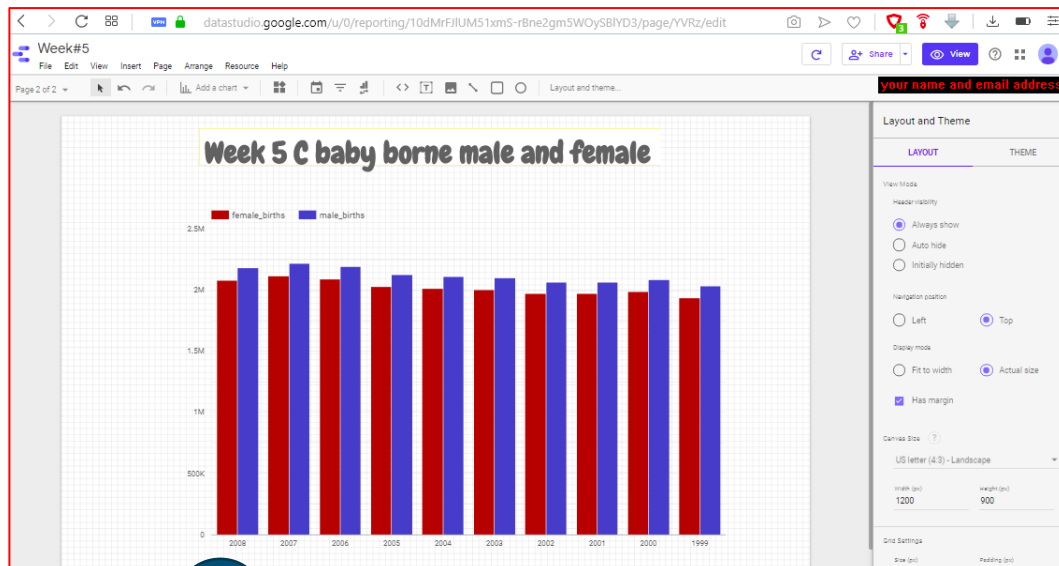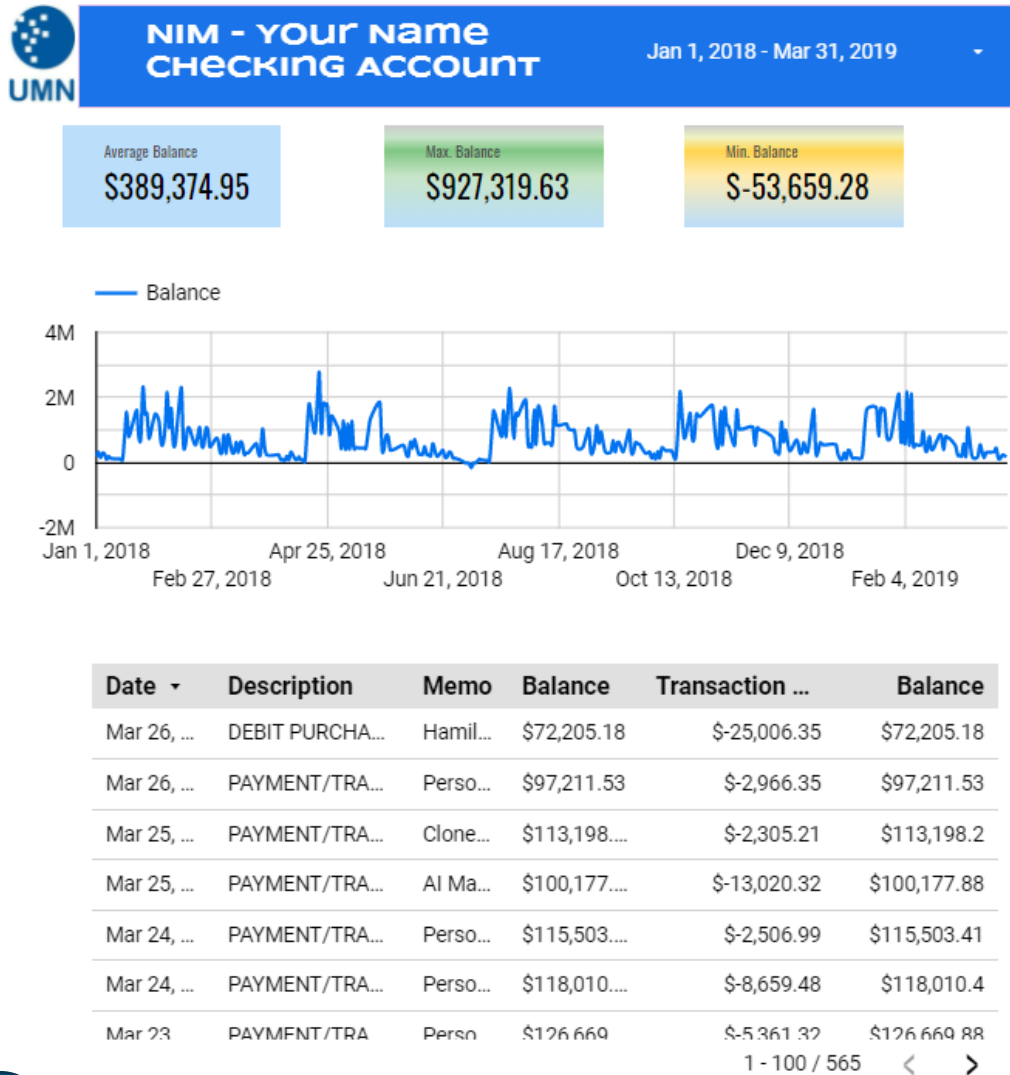
OUTPUT/ RESULTS
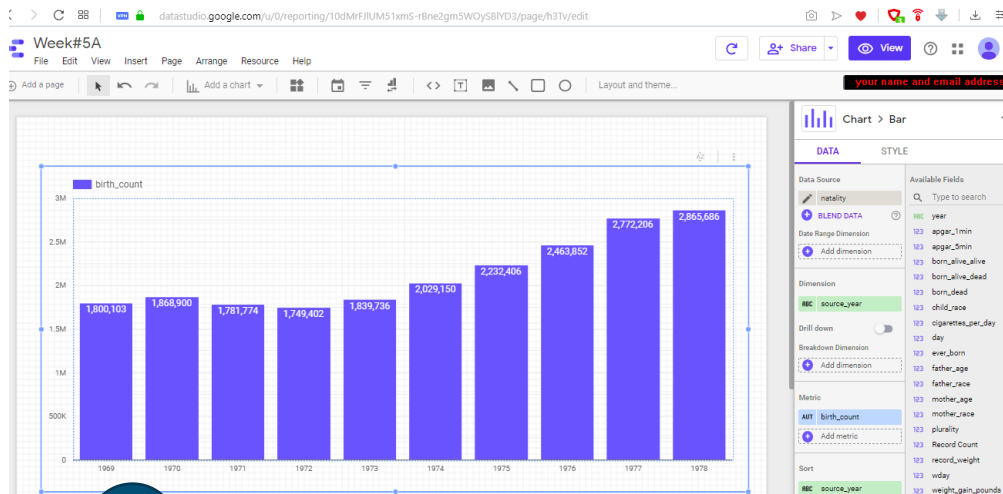
Figure 3.1 Birth Count by Year
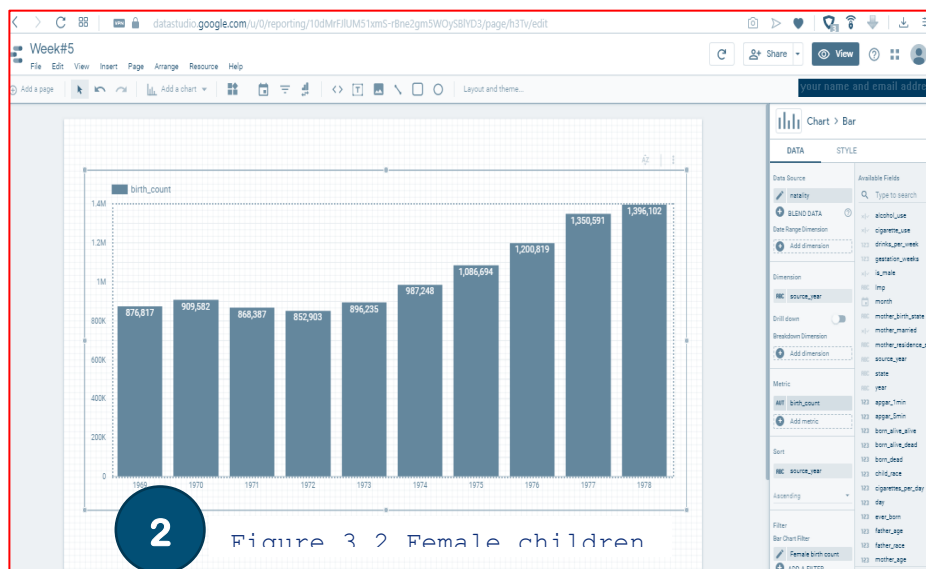
Figure 3.2 Female children

Figure 3.3 Baby borne Male & Female

IS-429 Week#3 Checking Account your_name

NIM - YOUR NAME CHECKING ACCOUNT
Jan 1, 2018 - Mar 31, 2019

Average Balance	Max. Balance	Min. Balance
$389,374.95	$927,319.63	$-53,659.28

Date ▾	Description	Memo	Balance	Transaction ...	Balance
Mar 26, ...	DEBIT PURCHA...	Hamil...	$72,205.18	$-25,006.35	$72,205.18
Mar 26, ...	PAYMENT/TRA...	Perso...	$97,211.53	$-2,966.35	$97,211.53
Mar 25, ...	PAYMENT/TRA...	Clone...	$113,198....	$-2,305.21	$113,198.2
Mar 25, ...	PAYMENT/TRA...	AI Ma...	$100,177....	$-13,020.32	$100,177.88
Mar 24, ...	PAYMENT/TRA...	Perso...	$115,503....	$-2,506.99	$115,503.41
Mar 24, ...	PAYMENT/TRA...	Perso...	$118,010....	$-8,659.48	$118,010.4
Mar 23	PAYMENT/TRA	Perso	$126,669	$-5,361.32	$126,669.88

1 - 100 / 565 ‹ ›

4 Figure 3.4 Visualizing a Bank Account

Data Last Updated: 2/8/2022 8:58:40 PM | Privacy Policy

REFERENCE

1. Documentation | Google Cloud
2. Cloud Google | BigQuery
3. Data Studio | BI Engine | Google Cloud
4. Other additional references are excerpts from various Online Learning/websites.

The End