Data Visualization in PowerBI

Mohamed Kassem

2021-04-27

Contents

1 Intorduction **5**

2 Setting and Objectives **7**

2.1 Setting 7

2.2 Objectives 7

3 Data **9**

4 Dataset **11**

4.1 Step 1 12

4.2 Step 2 12

4.3 Step 3 13

4.4 Step 5 14

4.5 Step 6 14

4.6 Step 7 15

5 DataModel **17**

5.1 Step 1 17

5.2 Step 2 18

5.3 Step 3 18

6 Report page **21**

6.1 information 21

6.2 Size 21

6.3 Background 21
6.4 Alignment 21
6.5 Wallpaper 21

7 Guide **23**

Chapter 1

Intorduction

This course is intended for educational purposes & preparing Data analysts for Exam DA-100: Analyzing Data with Microsoft Power BI : **Creating reports**

- All the needed resources to follow along can be found here https://github.com/Mkassem16/NycTaxiPBI.
- Creating basic PowerBI reports knowledge is prerequisite for this course.

By the end of this course Data Analysts should be able to:
+ Customize Report pages.
+ Decide for appropriate visualizations type.
+ To utilize PowerBI native visualizations.
+ Format and configure PowerBI native visuals.
+ Configure conditional formatting.
+ Import their own custom R or Python Visuals.
+ Utilize Slicers and Filters.
+ Adjust visuals for accessibility.
+ Configure automatic page refresh.
+ Create paginated reports.

Chapter 2

Setting and Objectives

2.1 Setting

The New York City Taxi and Limousine Commission (TLC), created in 1971, is the agency responsible for licensing and regulating New York City's Medallion (Yellow) taxi cabs. Over 200,000 TLC licensees complete approximately 1,000,000 trips each day.

For the sake of educational purposes Data Analysts should think of an Imaginary situation where NYC TLC has asked them to prepare a PowerBI report for their c-suite executives and decision managers to better under stand their monthly operations in general and track their daily and monthly KPIs.

2.2 Objectives

2.2.1 General

- Creating an easy to read and understand descriptive dashboard.
- Quick one look dynamic status updates.
- Tracking daily and weekly KPIs.
- Understanding revenue breakdown.

2.2.2 Requested

- Is there a specific day of week that NYC TLC should deploy more licensees?
- Is there a specific time of day that NYC TLC should deploy more licensees?

2.2.3 Extra

- Mining for relations between trip duration, distance and fare.

Chapter 3

Data

Data is publicly available at https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page

Please Find Data catalog explaining columns and categories here: https://github.com/Mkassem16/NycTaxiPBI/blob/main/data/data_dictionary_trip_records_yellow.pdf

For the purpose of building a demonstration dashboard, I extracted a sample of 10,000 rows of 2020 data. This sample dataset is hosted publicly on Google Cloud storage and can be queried directly from PowerBi. https://storage.googleapis.com/powerbi_datacamp/nyc_taxi_db.csv

```
glimpse(df)
```

```
## Rows: 10,018
## Columns: 17
## $ vendor_id           <dbl> 1, 1, 2, 1, 2, 1, 4, 2, 1, 2, 1, 2, 2, 2, 2, 2,...
## $ pickup_datetime     <dttm> 2021-03-27 12:47:16, 2021-06-10 19:02:02, 2021...
## $ dropoff_datetime    <dttm> 2021-03-27 13:39:54, 2021-06-10 19:31:53, 2021...
## $ passenger_count     <dbl> 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 2,...
## $ trip_distance       <dbl> 2.70, 15.10, 7.92, 6.50, 6.44, 10.00, 7.24, 36....
## $ rate_code           <dbl> 1, 1, 1, 1, 1, 1, 1, 5, 1, 1, 1, 1, 1, 1, 1, 1,...
## $ store_and_fwd_flag  <chr> "N", "N", "N", "N", "N", "N", "N", "N", "N", "N...
## $ payment_type        <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...
## $ fare_amount         <dbl> 29.0, 42.0, 26.0, 22.5, 24.5, 29.0, 22.0, 83.5,...
## $ extra               <dbl> 0.0, 0.0, 0.5, 0.5, 0.5, 0.5, 0.5, 0.0, 0.0, 0....
## $ mta_tax             <dbl> 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.0, 0.5, 0....
## $ tip_amount          <dbl> 5.95, 12.10, 5.46, 4.75, 3.87, 6.05, 4.66, 21.5...
## $ tolls_amount        <dbl> 0.00, 5.76, 0.00, 0.00, 0.00, 0.00, 0.00, 23.76...
## $ imp_surcharge       <dbl> 0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0.3, 0....
```

```
## $ total_amount        <dbl> 35.75, 60.66, 32.76, 28.55, 29.67, 36.35, 27.96...
## $ pickup_location_id  <dbl> 68, 138, 261, 262, 261, 100, 7, 132, 264, 170, ...
## $ dropoff_location_id <dbl> 162, 88, 41, 231, 162, 127, 53, 265, 264, 236, ...
```

Chapter 4

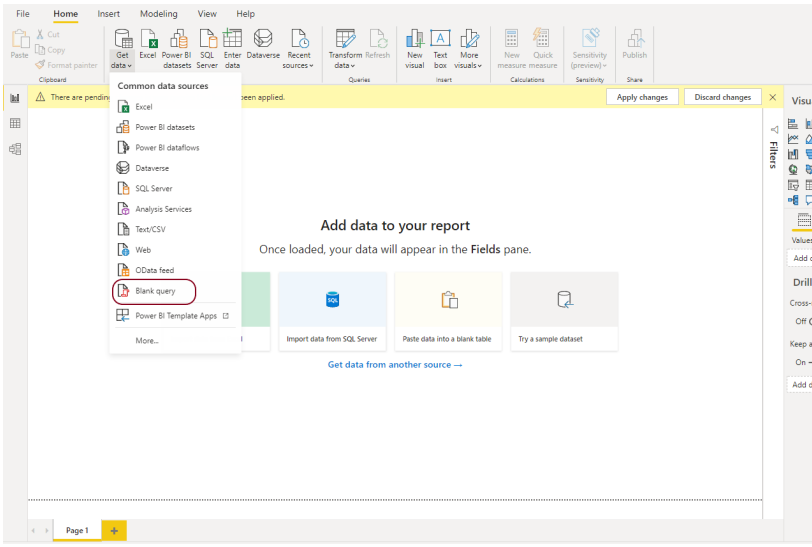
Dataset

Dataset in PowerBI vocabulary is: the set of queries that the user will perform in order to import all of his data tables and all the transformations that will take place before data can be ready for the model and visualization.

We will perform 2 queries to create 2 data tables:

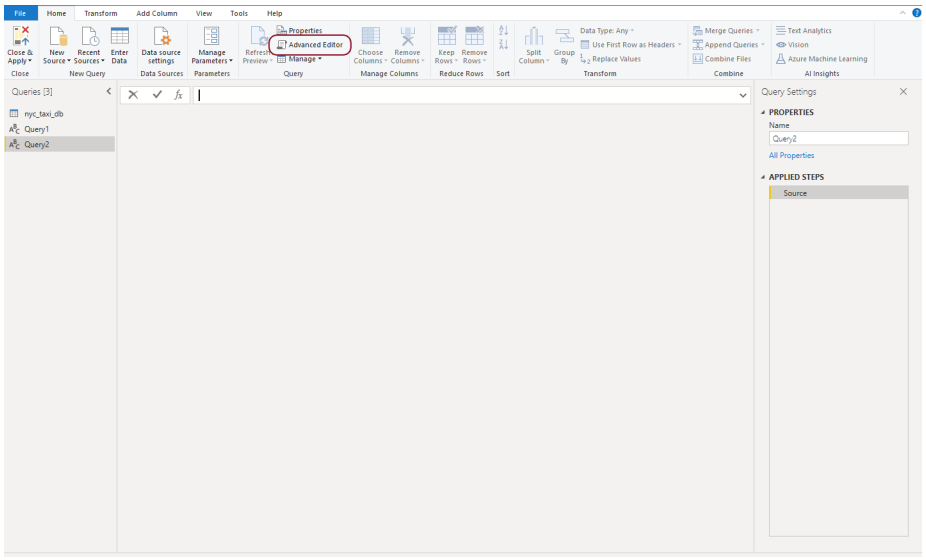
- Table 1 (**nyc_taxi_21**) : will include NYC TLC data. https://github.com/Mkassem16/NycTaxiPBI/blob/main/queries/nyc_taxi_21.pq
- Table 2 (**Calendar**) : Calendar table that will control the date dimension https://github.com/Mkassem16/NycTaxiPBI/blob/main/queries/Calendar.pq

4.1 Step 1



Getting data through Blank query

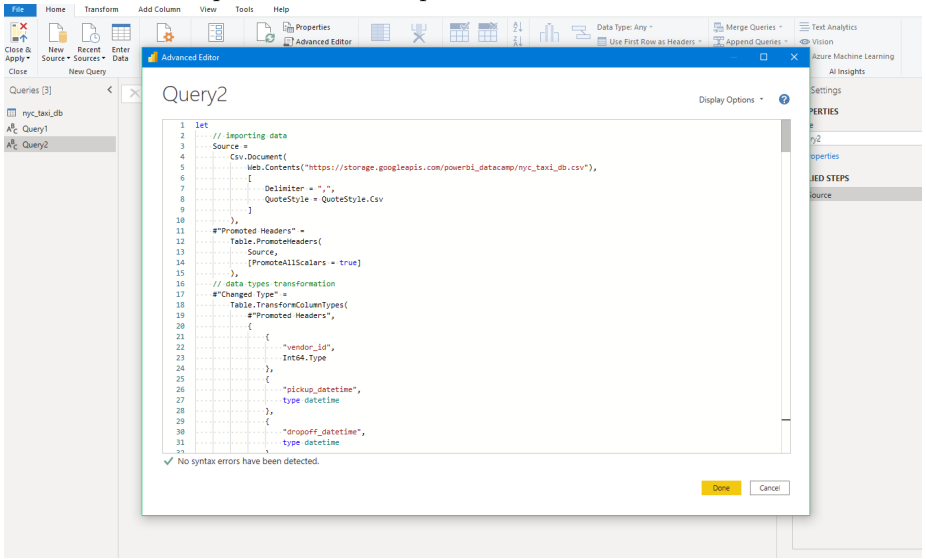
4.2 Step 2



Choose Advanced editor

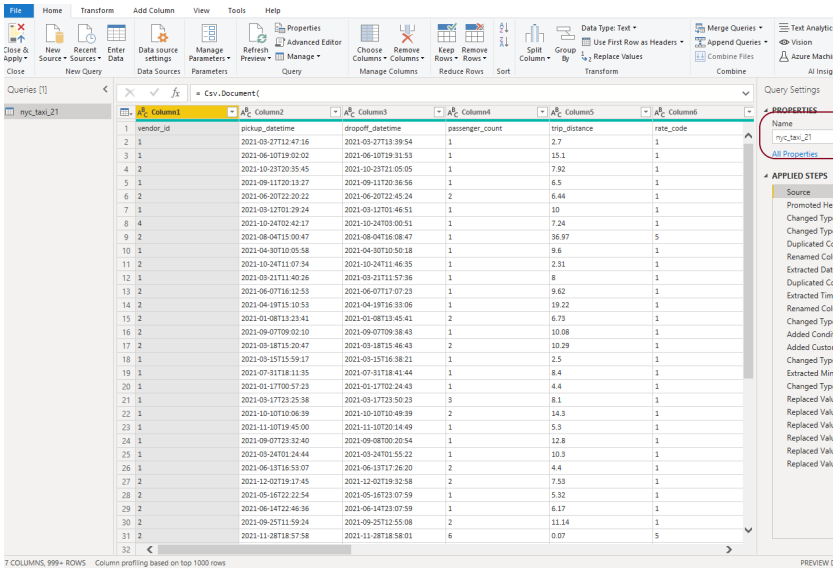
4.3 Step 3

Get the query of Table 1 from the provided link and paste it here instead of what



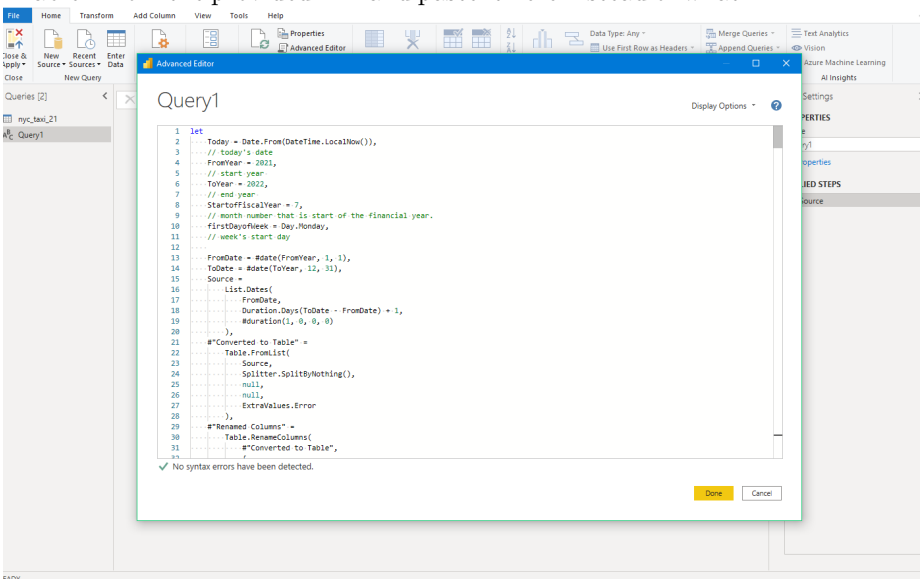
is already there
Step 4

Make sure that query name is **nyc__taxi__21**



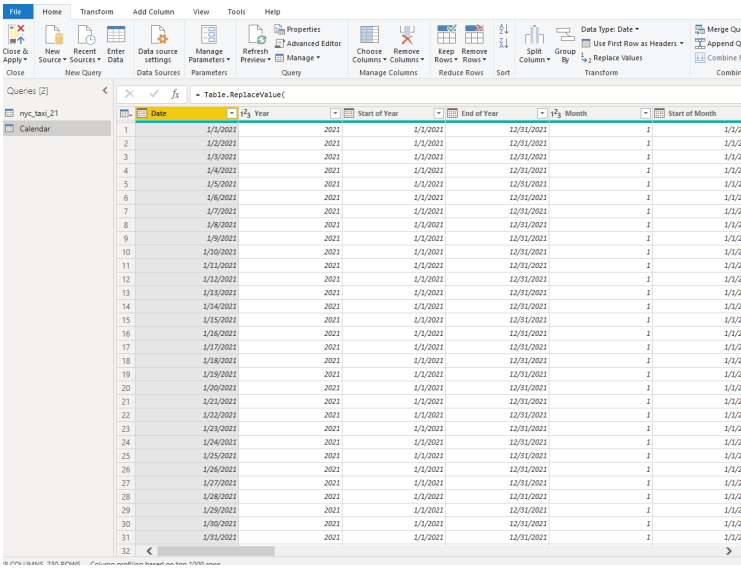
4.4 Step 5

Get the query of Table 2 from the provided link and paste it here instead of what is already there



4.5 Step 6

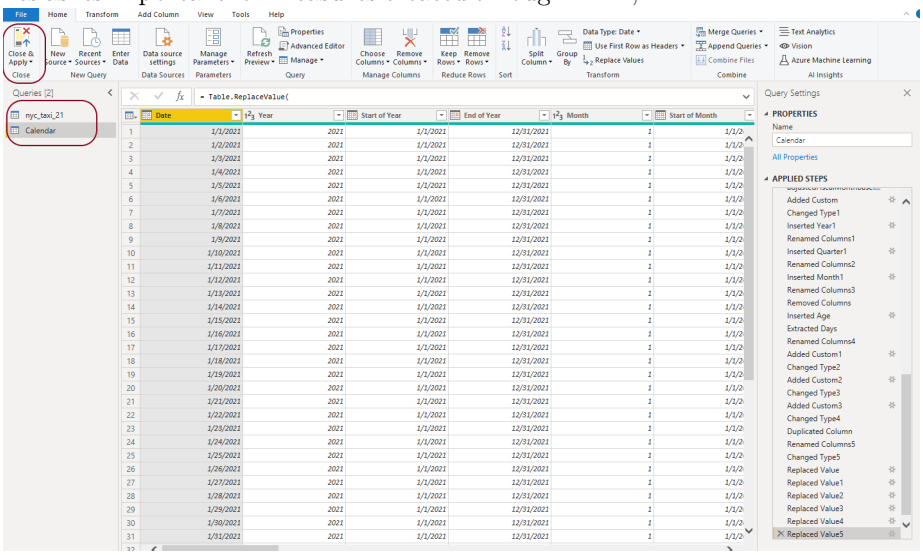
Make sure that query name is **Calendar**



4.6 Step 7

Double check queries names as its important for Measures created through DAX,

Then click close & apply



Chapter 5

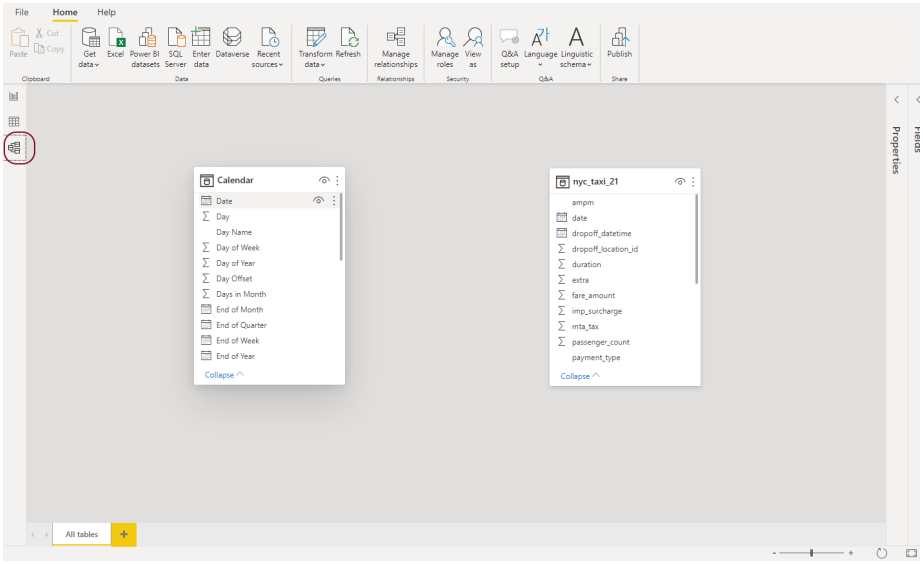
DataModel

Data Modeling in PowerBI vocabulary is: the step of creating relations between tables before proceeding and building slicers and filters on our visuals we need to make sure that the relationships are with the right direction, cardinality and that they are active.

We will create 1 relation between Calendar table and the data table

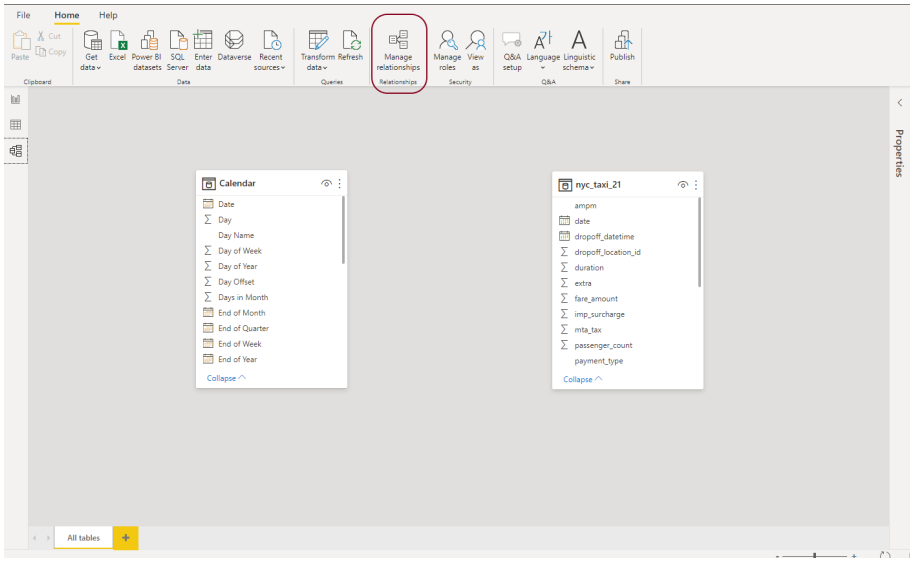
5.1 Step 1

Navigate to data modeling view



5.2 Step 2

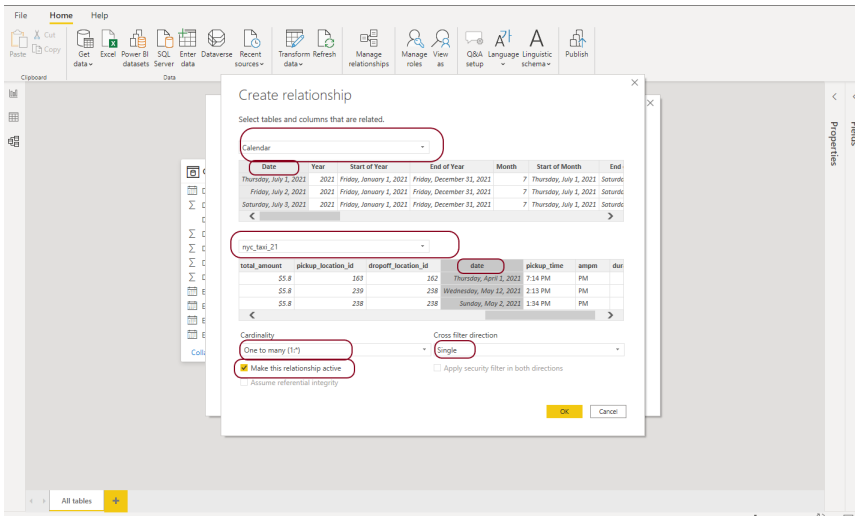
Click on manage relations



5.3 Step 3

Create new relation and make sure of the following:

- Select the right tables
- Select the right columns (by clicking on the column name in the display)
- Select the right Cardinality (usually with Calendar tables its One to many)
- Select the right direction (**Calendar** table to filter **nyc_tax_21** table)


File
Home
Help
Create relationship
Select tables and columns that are related.
Calendar
Date
Year
Start of Year
End of Year
Month
Start of Month
nyc_taxi_21
total_amount
pickup_location_id
dropoff_location_id
date
pickup_time
ampm
Cardinality
One to many (1:*)
Cross filter direction
Single
Make this relationship active
Assume referential integrity
Apply security filter in both directions
OK
Cancel
Properties
All tables

Chapter 6

Report page

We can start customizing report page by clicking on the format icon of the Visualization pane without choosing any visual.

Options include:

6.1 information

6.2 Size

6.3 Background

6.4 Alignment

6.5 Wallpaper

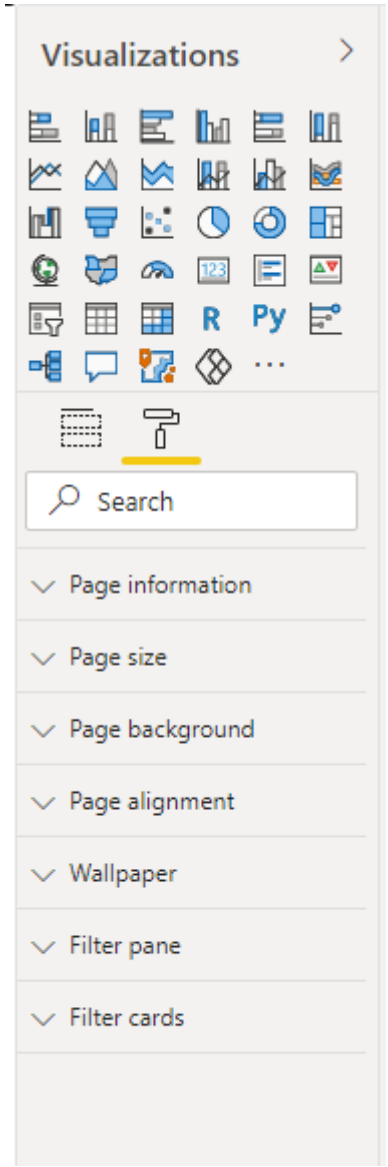


Figure 6.1: format

Chapter 7

Guide

Figure

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

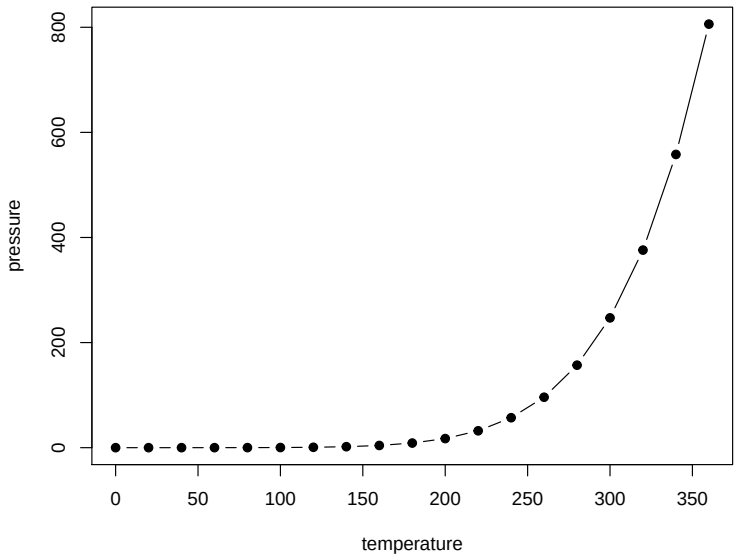


Figure 7.1: Here is a nice figure!

Table

Table 7.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

reference Chapter **??**

R

```
install.packages("bookdown")
# or the development version
# devtools::install_github("rstudio/bookdown")
```

link https://yihui.name/tinytex/.

bold **bookdown**

inclined *sample*