1) What's the number of relationships?  
- Try to get at least 2, based on both dataset. The scenarios are to be built based on the relationships you find.

2) Can I use other visualizations than those in notes and tutorial?  
- Highly encouraged to do so. You may find other ggplot2 or other plot types on your search engine.

3) How many graphs should we have?  
- Keep it to a maximum of 10.

4) Is there a word count?  
- No, however you should keep it concise. Visualizations play a big role, where visuals you produce should be self-explanatory, or can be explained in 1-2 lines.

5) Do we need to submit any presentation or Word document?  
- No, do everything in Google Colab and just submit the link.

This is a group assignment. Number of group members: 3 (preferred), 2 (minimum). Kindly inform your tutors on your group selection by end of Week 2. Members should be within same tutorial group. Ensure diversity in group members, meeting 1 of the following criteria: gender, nationality, race, native language. For instance, a group can have members from different genders, or different nationality.

**Goal**: Visualize the data to best represent the content, making use of both data sets.

**Data set A**: Government School Pupils by District 2017-2018  
Data rows: 1740  
Variables:  
1. School stage  
2. State  
3. District Education Office  
4. Year  
5. School type  
6. Sex  
7. Number of pupils

**Data set B**: Government School Teachers by District 2017-2018  
Data rows: 1748  
Variables:  
1. Year  
2. School stage  
3. School type  
4. State  
5. District Education Office  
6. Sex  
7. Number of teachers

**Tools:** R using Google Colab. You will need to work collaboratively between team members, as marks will be counted based on amount of collaboration.

**Task instructions:**  
Load both data files (use [sample Google Colab notebook](https://colab.research.google.com/drive/1v7wiWS8SRw4UqBuYkz0UgIBcvVhzMrg4?usp=sharing) to get started)  
As columns in both datasets do not align exactly, you will need to align them  
As the data provided is on school pupils and teachers, think of a relationship between them  
Based on the relationship you get from the data, use tools such as aggregation or differences (not limited to, but just examples), as data is given for year 2017 and 2018  
Based on the new variables, you will then think about how you want to present the data  
You may present it in combination of charts, tables, or graphs, based on what you get from the data  
Include references (URLs) for packages used, and if other similar work that is being done

**Submission instructions:**  
Only submission of the completed Google Colab link is required (ensure your respective tutor and lecturers are included as Commentor)  
Late submission: a penalty of 10% per day applies for each day of late submission or edited submission that is past due  
Do not make changes once the assignment is past due. Penalty applies strictly based on the last modified (submission) date and t