# LaTeX Template for SBE304 Project Proposal

Ali Gamal

firstauthor@eng-st.cu.edu.eg

Khaled Maher

secondauthor@eng-st.cu.edu.eg

Nada Ashraf

thirdauthor@eng-st.cu.edu.eg

- This template is based on the CVPR conference template[1].

- The information in this template is very minimal, and this file should serve you as a framework for writing your proposal. You may prefer to use a more collaboration-friendly tool while drafting the report with your class mates before you prepare the final report for submission. Remember that you should **submit both the report and code** you used for this project via Canvas. Also, **only one member per team** needs to submit the project material.

- The project proposal is a 2-4 page document excluding references[2].

- You are encouraged (not required) to use 1-2 figures to illustrate technical concepts.

- The proposal must be formatted and submitted as a PDF document on Canvas (the submission deadline will be later announced via the schedule & email).

- Please check out the text in these sections for further information.

## 1. Introduction

In this section, describe what you are planning to do. Also, briefly describe related work.

When discussing related work, do not forget to include appropriate references. This is an example of a citation [?]. To format the citations properly, put the corresponding references into the bibliography.bib file. You can obtain BibTeX-formatted references for the "bib" file from Google Scholar (https://scholar.google.com), for example, by clicking on the double-quote character under a citation and then selecting "BibTeX" as shown in Figure 1 and Figure 2.

---

[1] http://statcourse2018.thecvf.com/submission/main_conference/author_guidelines

[2] This means, references should of course be included but do not count towards the page limit
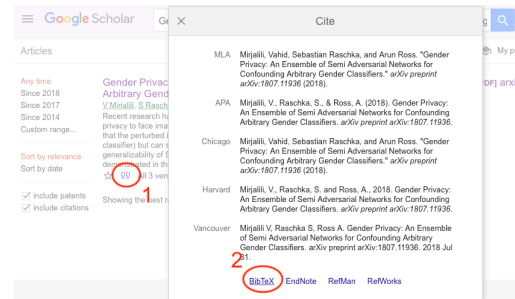


Figure 1. Example illustrating how to get BibTeX references from Google Scholar as a 1-column figure.

## 2. Motivation

Data science is an absorbing domain. Transforming raw data into meaningful insights is just magic. And when the data is biomedical, the task gets more interesting. We have great enthusiasm for this project because it meets our research interests and will provide us a great opportunity to learn new skills and dive into the awesome world of data science. Moreover, the data we will work on is collected from Egyptian patients and this relevancy makes us even more enthusiastic.

## 3. Problem Statement

## 4. Resources

What resources are you going to use (datasets, computer hardware, computational tools, etc.)?

## 5. Methodology

### 5.1. Exploratory Data Analysis

#### 5.1.1 Data variation

Exploring patterns of variation, typical values and outliers is an important task. We can gain such knowledge by visualizing the variables' distributions. To examine the distribution
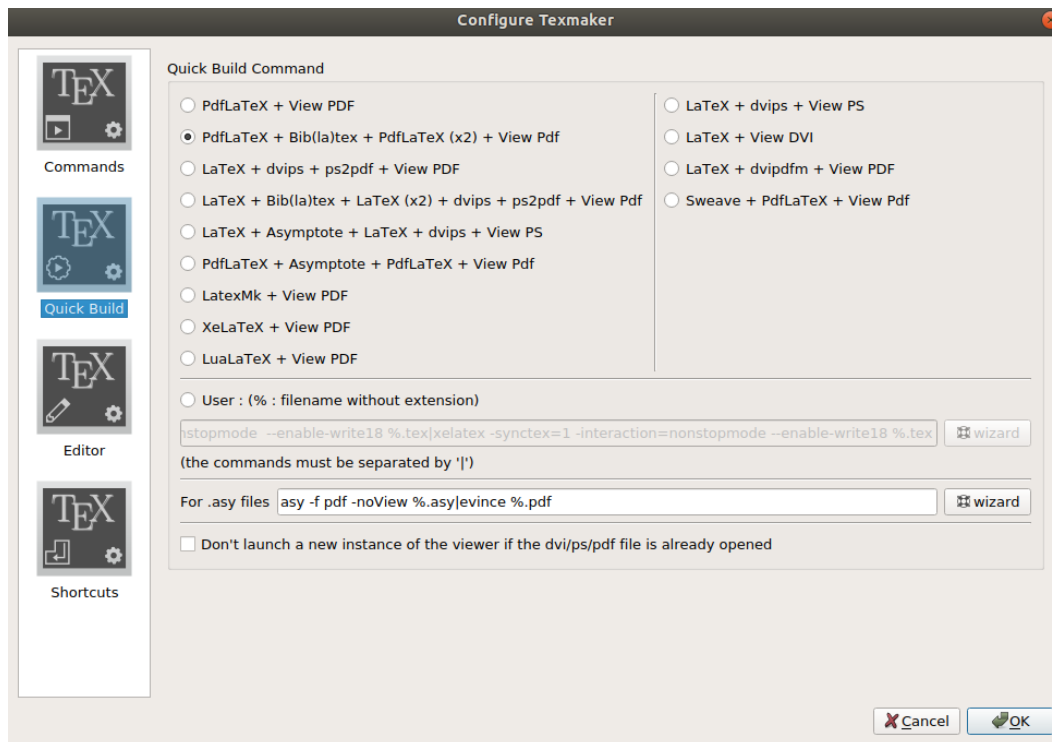
Figure 2. Compiling this document using *TexMaker*.

of a categorical variable, we can use a bar chart. And for continuous variables, histograms and frequency polygons can be used. To overcome binning bias of histogram and display all data, we can use swarm plots.

### 5.1.2　Co-variance

It's important to study the behavior between variables to gain useful insights that can be useful for feature selection. To examine the covariance between categorical and continuous variables we can use boxplot or violin plot. If both variables we are interested in are categorical we can use heatmap or scatterplot. If both are continuous, heatmaps can be used.

## 5.2. Data Preparation

### 5.2.1　Feature scaling

Using a normalization technique (Z-score or min-max normalization) to avoid skew towards high magnitude features.

## 5.3. Modeling

## 5.4. Evaluation

What would the successful outcome of your project look like? In other words, under which circumstances would you consider your project to be "successful?"

How do you measure success, specific to this project, from a technical standpoint?

## 5.5. Deployment

## 6. Milestones and Contributions

| Milestone | Date | Contributer |
|---|---|---|
| *EDA* | 29 Oct - 2 Nov | |
| *Pre-Processing* | 3 - 8 Nov | |
| *Model Building* | 9 - 15 Nov | |
| *Model Evaluation* | 16 - 18 Nov | |
| *Model Improvement* | 1 - 7 Dec | |
| *Deployment* | 8 - 10 Dec | |
| *Documentation and Publicity* | 11 - 13 Dec | |

## References