Introduction Writing

Think of the introduction of your paper as a hook. You need to use your first few lines as a lure to catch the attention of your readers. If readers don't find your hook captivating, you're not going to keep their attention. You know who your readers will be, you know why you're writing this paper, but you must tie together an introduction to start your paper off strong. This leaves some questions unanswered; how exactly *do* you write your introduction in a way that's appealing to the people reading your paper? By the end of this, you will have a great start to your introduction section, which will consist of 2 paragraphs:

1. The first paragraph should describe the **problem statement** and the **motivation** of the project.
2. The second paragraph should **describe the data** and its potential **limitations**.

The Purpose of an Introduction

An introduction should consist of a general outline of the paper and what you will be discussing. You should address the topic of the paper and any subjects related to this topic that you will be discussing or exploring. Make sure you include enough context about your topic in your introduction so your reader can have a general understanding of what argument you are going to make and the main points you will be discussing. You want to keep your reader invested in what you're discussing in the paper and show your reader why what you're discussing matters. You don't need to include actual evidence to support your points in your introduction, but you want to give your reader a sense that you do have evidence that supports your claim later in the paper. However, you don't want to go too in-depth with your descriptions – that is what the rest of the paper is for. One way to think about an introduction is that it serves as a plan or a roadmap that shows the reader where you will be going later in the paper.

The Contents of Your Introduction

For this project specifically, you will have 2 paragraphs in your introduction, as mentioned above. Before you begin writing your introduction, you should have selected a data set to work with.

First Paragraph:

In the first paragraph of your introduction, you will tell the reader what the **topic** of your paper is. To introduce your topic, give a little bit of background about your dataset's variables (you'll be explaining the dataset's other features more in depth in the second paragraph of your introduction). Along with the explanation of what you are looking at with your data set, you'll come up with **specific questions that you are trying to answer** about the data set that you have selected.

Coming up with the questions you want to ask may be tricky at first. To help you determine the questions you want to ask, think about the variables in the dataset you're working with. Answer these questions to help you determine what questions you'd like to ask and answer with your paper:

- What variables look the most appealing?
    - Why are they appealing to you?
- Are there possible relationships between variables that you'd like to look into?
- Is there a variable you think can predict a different variable?
- Can you look at how a variable changes over time?
- What questions will comparing these variables answer?
    - Why is that important?

These are some ideas that might help you get started on coming up with questions, but you are not restricted to them – be creative!

Once you have determined what questions you think you want to answer for your project, you can move onto the **problem statement**. Your problem statement is essentially changing your questions into statements that describe the questions you will answer. If you have written papers for other classes before, you can think of this as a **thesis statement**. Below are a couple of examples on how you could do this:

<u>Example 1:</u>

Question: What is the relationship between the weight of a car and its gas mileage?

Problem Statement: The relationship between the weight of a car and its gas mileage will be investigated.

<u>Example 2:</u>

Question: How does the amount of sunlight that a plant receives affect its growth?

Problem Statement: The impact of sunlight on plant growth will be analyzed.

You'll then incorporate the problem statements into your introduction paragraph to show readers what you're investigating with the dataset. This will help your readers understand what you will be answering by the end of the paper.

Lastly, your first paragraph should include your **motivation** for your project. This is where you tell the reader why you are choosing to investigate the questions you decided on as well as why they should care about the answers to the questions. Think about why it's important that your questions are answered. Answer these questions to help you understand your motivation behind your study:

- Why are these the questions I am asking about this dataset?
- Who could this impact or benefit?
- Why do this study's results matter to my reader?
- What will someone be able to do with this information?

Take the time to explain why you are investigating these questions and this dataset to your readers, so they understand the importance of your study.

To summarize, your first paragraph should include:

- A short discussion of your topic
- The questions you will answer
    - Written as problem statements
- Your motivation for the project

Once you think you are done writing the first paragraph of your introduction, check it against the requirements in the rubric to make sure you included everything:

Project Goals and Motivation			
60	**51**	**39**	**30**
Questions are explained in detail and motivation is well established for why the reader should care about the questions and their answers.	Questions are explained enough, and fair motivation is given for why the reader should care about the questions and their answers.	Questions are stated but motivation lacks reasoning for why the reader should care about the questions and their answers.	Questions are not clear, and no motivation given.

Second Paragraph:

In the second paragraph of your introduction, you will be describing your data set to the reader. This includes identifying your cases, target population, and sample and discussing the variables you will be using.

First, here is a reminder of what cases, target population, and sample are:

- **Cases**: individual subjects or observations
- **Target population**: the entire group that you want to make conclusions about
- **Samples**: a subset of the entire population – used to represent the whole population

You will discuss each of these in your paragraph. You should briefly discuss the population that you are hoping to learn about through your project – this is your target population. You should also discuss the sample from the population that you are using for your study. This could include information such as the size of the sample or the time period of the sample. For the cases, think about what the observations are in your sample. Are they people? Objects? Events?

Take the time to write out your description of the dataset by explaining the subjects, observations, target population, samples, and other relevant information in a way that will make sense to your reader.

Below are the requirements for this section of the second paragraph:

Data Description			
20	**17**	**13**	**10**
Cases, target population, and sample are correctly identified and stated.	Cases, target population, and sample are mostly correctly identified and stated. Mistakes are minor.	Cases, target population, and sample are somewhat correctly identified and stated. Mistakes are significant.	Cases, target population, and sample are not correctly identified or stated.

In order to determine which variables you want to use in your project, think about the questions you decided on in the first paragraph – what variables are needed to answer these questions? These are the variables that you should be using, and thus these are the variables you should describe to your reader.

When discussing your variables, you should give a brief description of what they are (unless it's obvious, such as if the variable is 'year'), as well as the type of variable. Possible variable types are:

- Discrete
    - Numerical
    - When there are a finite/countable number of values
        - Often integers
- Continuous
    - Numerical
    - Can take any value in a given range
- Nominal
    - Categorical
    - No inherent order/ranking of groups
- Ordinal
    - Categorical
    - Groups have a meaningful order/ranking

You should also identify whether the variables are explanatory or response variables if your question focuses on the relationship between two variables. The explanatory variable would be the variable you believe affects the other variable, and the response variable is the variable you think is affected by the explanatory variable.

Take the time to write out your findings about the variables. Think about these questions:

- What is every variable in the data set and its type?
    - What is the variable giving information about?
    - Are they discrete, continuous, nominal, or ordinal?
    - Which are the explanatory variables that will affect the other variable?
    - Which are the response variables that are affected by the explanatory variables?
- Which variables will you specifically be working with to answer your questions?
- Is there a relationship between two variables? Explain what you think it will be.

Once you have done this, compare what you have written to the rubric requirements for this section:

30	26	20	15
Variables are well described, and their types (discrete, continuous, nominal, or ordinal) are correctly identified and stated. Explanatory and response variables are correctly identified and stated.	Variables are somewhat described, and their types are mostly correctly identified and stated. Explanatory and response variables are mostly correctly identified and stated.	Variables are stated, but their types are not correctly identified and stated. Explanatory and response variables are somewhat correctly identified and stated.	Variables are not stated, or their types are not identified. Explanatory and response variables are incorrectly identified or missing.

Lastly, you need to include a discussion of your **data source**. This includes addressing the questions below, as well as any other information you believe is important for the reader to know.

- Where did your data set come from?
- Who collected the data?
- How was the data collected?
- What was the purpose of collecting this data?

You also want to address any potential biases or limitations that your data set might have. Based on your answers to the questions above, could there be any biases in your data? For example, was the data collected in a way that could make it biased or potentially non-representative of the population? Think about other ways that the data could be limited or restricted. Some examples of things that could potentially limit or restrict your data are if your sample size is small, if the data is from many years ago (meaning it could be outdated), or it is limited to a certain geographical location (meaning it might only be representative of that population and not a wider population).

To address biases and limitations, answer the following questions:

- Are there any limits or restrictions on the data?
    - Is it limited to a certain amount of people?
    - Is it from many years ago (possibly outdated)?
    - Is it limited to a certain geographical location?
    - Is it restricted to only a certain population?
- Do these limitations/restrictions cause any bias?
- Does the collection method raise any other questions about bias?
    - Was it collected in a way that prohibited certain people from participating?
- Are there any other reasons your data may be biased or limited?

Once you have written about your data source, check with the rubric below to make sure you met all of the requirements:

25	22	17	12
The data source is stated, collection method is described, and biases and restrictions are well explained.	The data source is stated, collection method is somewhat described, and biases and restrictions are somewhat explained.	The data source is stated, but collection method or biases or restrictions is missing.	Data source is stated but collection description, biases, and restrictions are all missing.

To summarize, your second paragraph should include:

- Cases, target population, and sample
- Variable descriptions
- Variable types
- Data source description
- Any potential biases or restrictions in the dataset

At this point, you should have completed a rough draft of your introduction paragraphs for your project. However, keep in mind that this is only a rough draft. As you continue working on your project, you might find that some of the information you have written will need to be changed – that is okay! You should now be more familiar with your data set and have a good understanding of the questions that you want to answer. This is a good start for your project and will be very useful as you move into the next sections of the project.

If you want to go over anything you have written or you want help brainstorming or getting started, feel free to book an appointment with the Writing Center. We would be happy to help you! You can scan the QR code below or visit simpson.mywconline.com to book an appointment.