

## Response Summary:

### 1. Student Information \*

<b>First Name</b>	Grace
<b>Last Name</b>	Combs
<b>Major</b>	Web Development
<b>Course</b> (e.g. CGT 270-001)	CGT 270-003
<b>Term</b> (e.g. F2019)	SP2022

### 2. Email Address \*

(University Email Address is required.)

gcombs@purdue.edu

### 3. Visualization Assignment \*

- Homework Assignment

## Generate

### 4. Identify appropriate data sources: is the data publicly available? What search methods were used? \*

<b>Data source 1</b>	<a href="https://nces.ed.gov/programs/digest/d19/tables/dt19_209.10.asp">https://nces.ed.gov/programs/digest/d19/tables/dt19_209.10.asp</a> This is government data and is publicly accessible. This data was collected using a head-count of full time and part time teachers excluding pre-k teachers.
<b>Data source 2</b>	N/A
<b>Data source 3</b>	N/A

### 5. Data format: what format is the data in? Structured vs instructed? All text, a combination, multiple sources? Is it primary or secondary data? \*

The data is organized and structured. It has multiple government sources. This data seems to be secondary data because it was collected from multiple sources and compiled together.

### 6. Data types: what types of data are in the data? How are they stored? What is the access to the data (API, JSON, txt, csv, etc.)? What structure holds the data (data base, spreadsheet, etc.)? \*

The data is accessible online and also in .xls format. The data is formatted using a spreadsheet however the spreadsheet is customized and not exactly like a typical excel spreadsheet.

## Evaluate

### 7. Variables: list the data variables? What are the parameters? Give them names. What are the dependent variables and independent variables? \*

There are a lot of variables here but most are characteristics of the teachers. The dependent variables are the number of teachers and the independent variables are the individual characteristics.

**8. Audience & Assumptions: list any assumptions you have about the data. Who is your audience? \***

I am assuming that this data will display the U.S. having more black teachers than were present in W.E.B Du Bois' data. My audience is the general public and people interested in race-based statistics.

# Generate

**9. What real life behavior does the data reflect? Does it show patterns of activity, regularity of events, a timeline, population data, etc? Explain. \***

It shows population data within the context of teachers only, and also shows a timeline of increase or decrease in black teachers.

**11. What are the weaknesses of the data source? Is it likely that the source will be available in the future? Is the data complete? What is the quality of the data? Is it specific to your needs for. the current project? Is the data in the format you need? Are there missing data? Explain. \***

The data is incomplete because each year more can be added. This source will likely be available in the future. The quality is high and meets my needs.

**12. What information is emphasized? What is the central focus of the data? Explain. \***

Number of teachers having certain characteristics.

**13. At what level of granularity is the data provided? Is the data summarized, or do you have access to the raw data? Is the data categorized or is the data in a format that allows you to create your own categories, etc. Explain. \***

The data is raw data and categorized.

**14. What is the scope of the data? What topics can be covered using the data? Is there a time range/frame? Is the data for a specific area/discipline/demographic etc.? Explain. \***

A lot of topics could be covered using this data, such as the gender and age of teachers. There is a time frame from 1987-2018

---