**Teaching Evaluation Data Project**

**Clean**

To begin working with this dataset, I compiled all of my teaching evaluations since moving back to the United States in 2018, which is when I joined the EAP program at The George Washington University as a Teaching Assistant Professor. To do this, I pulled down all annual data from the university’s teaching evaluation website and transposed it into an *Excel* worksheet. At first, it was just a direct move, capturing the data as it existed in its raw form in the original data set. This original transposition was done in a wide format, matching that procuded by the university’s system. From there, I set about cleaning the data.

My first step was to combine categories, as there was no small amount of overlap between certain categories. For example, the original surveys asked for an overall instructor rating and an assessment of the instructor overall. This was combined into a single new category: “Instructor Rating (Cleaned)”. Similar steps were taken with categories related to fairness in grading and timeliness of grading and feedback, and improvements in academic language use, with their child categories collapsing down into “Grading Fairness (Cleaned)” and “Timeliness of Grading and Feedback (Cleaned)”, and “Academic Language (Cleaned)” respectively. To collapse these down, an average of the two scores were take in cases where both questions were asked of the same group of students—which is often done as a sort of control to account for Likert question/survey fatigue.

Further review of the data showed no small number of null responses. These were left in place to account for the fact that not all questions were asked of all sections. In cases where a question has only been asked of a single section—such as in the case of either my Teacher Development or Professional Presentation course (both of which I have only taught a single section of so far)—these questions were dropped from further analysis, as they can be seen as not being statistically significant compared to the rest of the data.

Additionally, rating scales were normalized to an ordinal scale from 0.0 to 5.0 and grade level information was normalized to a nominal scale from 0-2, with 0 representing non-degree students, 1 representing undergraduate students, and 2 representing graduate students.

**Transform**

To begin transforming the data, I started with my guiding question, which was how my teaching assignments and evaluations have changed over time, and how can I effectively represent that to a third party—which in this case refers to university administrators. I began by looking at my actual teaching assignments and using the **=COUNTS()** and **=COUNTIFS()** functions in *Excel* to get a count of how many of each level of course I have taught between Summer 2018 and Summer 2022. This count included all classes taught at all institutions where I hold any appointment (full-time, teaching faculty at George Washington; part-time, workforce development faculty at Montgomery College). This also allowed me to quickly count which department my teaching had taken place in—Teacher Education & Development, English for Academic Purposes, Columbian College of Arts and Sciences, English as a Foreign Language, or the English Language Center.

To facilitate getting to know my data at a glance, a calculated row was added to each column in the enrollment, response rate, and evaluation data sections using *Excel*’s **=AVERAGES()** function. Additionally, a calculated column was added to divide number of students in each class by the number of evaluation responses in each class to arrive at the response rate. This was then formatted as a percentage using the **format cell** feature.

The data’s current wide format may not be the best format for all data analyses. To facilitate further analysis, the data was transformed into a long format as well. I used a variety of math functions to double check the numbers, including: **=SUM(), =SUMIF(), =COUNT, =COUNTIFS(), =AVERAGE(),** and **=AVERAGEIF()**. I then added a calculated row to get the average scores for each question by year, using the **=AVERAGE()** function. From there, I created a few helper tables to drive certain parts of the dashboard, which will likely be replaced in v2.0 (see below).

*For version 2.0*, I intend to include a worksheet where all data has been transformed into an ***Excel* Table** to make it easier to add data each year. I also intend to drive fewer of the dashboard features from transformed data or off of hardcoding, relying more on the use of pivot tables and pivot charts paired with slicers to do the heavy lifting and to create a more scalable report in the end that is better future proof. Additionally, v2.0 will include slicers to separate information by institution, as my role at Montgomery College persists.

**Explore**

To explore the data, I set about creating a series of initial aggregate charts to provide at-a-glance as well as deeper insights into the data. I created a basic aggregate chart showing evaluation scores between 2018 and 2022 (bar chart), as well as a chart comparing my aggregate average scores to those of all instructors in the Columbian College of Arts and Sciences (CCAS) (bar chart). I additionally created two pie charts to show the proportion of classes taught at each level and the response rate for teaching evaluations.

Additionally, I created a series of **pivot tables** to provide different summaries of the data based on my guiding questions, which sought to address how my teaching assignments and evaluations have changed over time.

*For version 2.0*, I intend to rely a great deal more on **pivot tables** instead of hard-coded or calculated helper tables. This should, when paired with the use of an ***Excel* table** for managing the aggregate data, help to further future-proof the workbook as a reporting tool.

**Analyze**

To analyze the data, I began by further altering my **pivot tables** to work off of averages instead of sums using the **pivot table options**. I also renamed row identifiers to be more descriptive, moving away from terminology like “Average of Instructor Rating” to “Instructor Rating”. For these pivot tables, I created multiple versions, with some using the evaluation questions as the rows and years as the columns, and some using the years as the rows and the evaluation questions as the columns.

It was at this point that I discovered my first error, perhaps due to my initial reliance on calculated values instead of pivot tables. Namely, there were numerous circular references that were causing problems. So, before continuing with the analysis, I used the **error checking** feature under the **formulas** **ribbon** to help quickly identify where these errors were occurring and to rectify them.

With that complete, actual analysis of the data could commence. Regarding the first question, of how my teaching assignments had changed, we see some interesting movement here. First, my primary role is most certainly as an undergraduate instructor, with the vase majority of my classes taking place at that level. Moreover, most of my teaching takes place in the EAP department (my primary, full-time appointment), and occurs during the fall semester. After that, most of my teaching is in the summer, which again is inline with my initial hiring appointment on a summer/fall rotation. Now, with the pandemic, enrollment patterns have changed, and we see that reflected in the data. Around 2020, we see a regular increase in my spring semester assignments. Another change that we see when it comes to appointments is the graduate addition of courses either outside of EAP, as I flex to new areas, or to new course levels (i.e., graduate and non-degree) as I take on new roles in my home department(s) to meet departmental need. This showcases considerable flexibility on my part as an instructor.

When it comes to my evaluations several interesting trends have emerged. The first is that my response rate for evaluations is relatively low, at 41% between 2018-2022. I do not, however, have data for how this compares to my departments or my parent colleges. What I do see in the data, however, is a steady decline in response rate beginning around 2020 after pandemic measures were instituted at the university, which included a move to online teaching. Looking at the overall data, as I compared to other instructors in the Columbian College of Arts and Sciences at GW,[[1]](#footnote-1) I am either performing inline with or slightly above the college average in almost all categories. Indeed, the only places where I appear to be underperforming is in my oral communication classes’ assessments. However, the case can be made that these should not be considered, as there simply isn’t enough data to make a useful comparison. I have only taught a single section of this course, compared to the dozens of sections of written communication that I have taught. That being said, it does underscore my relative newness to teaching these kinds of classes, and that may be reflected in the initial evaluation data. It would be interesting to see how those change over time. Looking at my individual data year-over-year, we do see a slight dip in overall performance across major categories in 2022, but that may be (at least in part) due to the partial nature of the 2022 data subset. It could also be connected to the sheer number of new classes that I have taught during 2022, as well as a new focus on skilling up in data and computer sciences to help facilitate a change in my professional focus going forward. Continued monitoring of this data will be important to ensure that overall teaching quality is staying consistent with other years, where I was again at or above the CCAS average.

**Visualize**

To begin the visualization (DataViz) process, I began by sketching out possible dashboard designs. During this time, my focus was on planning which elements I would need to make an effective dashboard, what layout options might best support the story that I want my data to tell, and how I wanted my user to interact with the data. This included thinking of the various interaction paths a user might take and how they would relate to different design elements. Once this was complete, I started creating my interactive dashboard and the charts that would make it up.

I began by creating a section that shows the number of students taught in a given year and the number of evaluations received. To power this, I used a **data slicer** to provide my user with selectable buttons that would update the counts in their respective fields. This allows a user to see my over instructional impact as well as how many evaluations I received in that year—to better inform how they weight that data.

I then created a section to show the proportion of classes taught in each department for which I have either primary or an affiliated faculty appointment, as well as the proportion of classes taught in each semester. To make these more compelling, I used **conditional formatting** and the **data bars** style to create cells that would not only show the counts for each variable, but that would fill the cell proportionately to the total. To help power this, I set my min value to 0 and used a **=SUM()** function for the max value to help future-proof this so that it would account for new data, auto updating both the data and the data bar. To help support this auto update, I also added **=COUNTIF()** functions to relevant areas of the data to ensure that this part of the dashboard, and its requisite helper tables, would properly stay up-to-date.

From there, I created a **pivot chart** from the pivot table mentioned above. This chart was designed to be a key interactive element, allowing the user to pick totals or individual years and to see my evaluation results for that selection. This, again, required the use of a data slicer to help drive the selection and update process. It also required some changes to the pivot table. Namely, I changed the sums to averages and renamed categories to be more user friendly, in keeping with the process outlined earlier in this report. I also removed all extraneous selection buttons from the chart, so that the user could focus on what I wanted to be the primary means of interacting with the data, as presented through the data slicer. To better reflect the actual data, I changed the scale of the axis from 4.5-5 to 0-5, with major breaks changing from .1 to .5 and minor breaks from .01 to .05. I also added in a chart title so that users knew what the chart represented at-a-glance, as well as simplified instructions for using the interactive aspect of the chart.

The final aspect of the DataViz was the unify the design. For this, I chose the university’s color scheme of “buff” (tan) and blue (navy blue). This was applied to the “Instructor Evaluation Scores” chart to quickly distinguish columns, to the “Response Rate” chart to distinguish responses vs. total possible respondents, and to the “Instructor/CCAS Evaluation Scores” chart to distinguish instructor and CCAS-average scores.

**Share**

This dashboard will be shared with relevant decision makers when I submit my reappointment dossier to GW in December. Given the high-stakes nature of this process, certain precautions where taken to protect the raw data driving the dashboard. First, I hid all of the data sheet from user view, as they have no need to access the raw data. To ensure that sheets were fully hidden from end users, I used the **VBA** option under the **Developer Tab**, setting all data sheets to **2-xlsheetveryhidden**. This means that hidden sheets will not trigger excel to make “unhide” available when right clicking on sheet tabs, acting as if there’s nothing hidden in the first place. To protect the structure of the workbook, and the relationships across worksheets, I password protected the entire workbook using the **Protect Workbook – Structure** options. I further protected the dashboard itself using the **Protect Worksheet** option, ensuring that users could still use the slicers and that charts/figures would still update appropriately.

To support those users unfamiliar with dynamic dashboards, both a set of written instructions and a short screen-capture video were generated to walk them through using the dashboard. This dashboard was then supplemented, as required by the university with a written summary of the data.

1. Comparison data for Montgomery College is not available to instructors [↑](#footnote-ref-1)