Unit 2: Lesson 14 - Creating Summary Tables

Background

A summary table is a table used to compute summary statistics of a larger dataset. Summary tables are another way that computational tools can be used to look more closely at data in order to identify trends and patterns. They can often be good data visualizations on their own, and they are quite useful when trying to make charts from larger collections of data.

**Overview**

In this lesson students learn how create their own summary tables from raw data. A summary table typically represents one or more aggregations (groupings of items) and computations that are performed on the raw dataset. In most spreadsheet programs, a summary table is called a pivot table. In the lesson, students learn how to make pivot tables in Google Sheets using a provided dataset. Then students turn to the data they’ve collected as a class and, with their partner, use pivot tables to investigate it further.

**Purpose**

Making a summary (pivot) table is often considered an advanced technique. Once you get used to it, however, it's an extremely powerful computational tool that is available in most spreadsheet software. The purpose here is to acquaint students with using such a tool and to expose this power. Also creating summary tables is a direct tie to the CSP Framework essential knowledge statement: 3.1.3C Summaries of data analyzed computationally can be effective in communicating insight and knowledge gained from digitally represented information.

The other purpose here is that creating a summary table is a good example of making a computational artifact for the Explore Performance Task. For that performance task students might find some raw data while doing research and might create a new artifact that is a summary table of the data that reveals some interesting aspect of it. Using a tool like a spreadsheet to make summary tables let's you explore data in deep ways, quickly and easily.

Being able to manipulate data is an important skill for computer scientists. Being able to create summary tables from larger datasets represents a form of computational thinking. To make a good summary table, one must have a good sense of the data, be able to hypothesize about what might be interesting to look at, and then have the skills to use a computational tool to create it. While seemingly mundane, a spreadsheet is an extremely powerful tool for working with data. Understanding the features of a spreadsheet tool, and what kinds of computations it can perform, can save you a lot of time and energy from either doing such things “by hand” or writing your own program to do it.

Vocabulary

* Summary Table: A table that summarizes information about some larger dataset. It typically consists of performing computations like sums, averages, and counts on higher level groupings of information. The intent is to summarize lots of data into a form that is more useful, and easier to "see".
* Pivot Table: The tool used by most spreadsheet programs to create a summary table.
* Aggregation: A computation in which rows from a data set are grouped together and used to compute a single value of more significant meaning or measurement. Common aggregations include: Average, Count, Sum, Max, Median, etc.
	+ For example, if some dataset contained information about how many hours of television people watched and included their age, you could "aggregate the data by age" and compute the average hours watched for each age group. You could also "aggregate by hours of TV watched" and compute the average age for each number of hours.

Lesson

* Learn how to make summary tables in Google Sheets.
* Make two (2) summary tables of your own data.

Resources

* All activities are presented here in Code Studio.
* Click Continue to move on.

**WATCH for pivot tables:**

**https://www.youtube.com/watch?v=IaOtCWxx3xU**

**http://www.excel-easy.com/data-analysis/pivot-tables.html**