

Quick Survey Please!

<https://bit.ly/3ekjPzD>



Twitter: [#OTFDataAnalysis](#)

[@pbeens](#)

[@grant_hutchison](#)



Addressing the Data Analysis Expectations: Using Google Colab to Analyze and Graph Your Data

Twitter: [#OTFDataAnalysis](https://twitter.com/OTFDataAnalysis)

Presentation: bit.ly/OTFDataAnalysis

Repository: bit.ly/3gZNPIN or is.gd/BeensOTFdata



Peter Beens & Grant Hutchison

[@pbeens](https://twitter.com/pbeens)

[@grant_hutchison](https://twitter.com/grant_hutchison)



1.2 determine the positions of individual data points in a one-variable data set using quartiles, percentiles, and z-scores, use the normal distribution to model suitable one-variable data sets, and recognize these processes as strategies for one-variable data analysis

2.4 determine, by performing a linear regression using technology, the equation of a line of best fit for a set of data, determine the fit of an individual data point to a linear model (e.g., by using residuals to identify outliers), and recognize these as strategies for two-variable data analysis

1.1 distinguish situations requiring one-variable and two-variable data analysis (e.g., associated numerical summary statistics, charts, summary tables) and bar graphs, and line graphs

A2.1 describe ways in which geographic investigation can help them develop skills, including spatial technology skills and the essential skills in the Ontario Skills Passport (e.g., reading text, including graphic text; writing; oral communication; using maps, graphs, charts, tables; computer use; use of geographic information systems [GIS], satellite imagery; measurement and calculation; data analysis; decision making; planning; organizing; finding information; problem solving), that can be transferred to the world of work and to everyday life

D2.4 demonstrate the Essential Skills that are important for success in the health care industry, as identified in the Ontario Skills Passport (e.g., computer use, data analysis, job task planning and organizing);

THE ONTARIO SKILLS PASSPORT: MAKING LEARNING RELEVANT AND BUILDING SKILLS

The Ontario Skills Passport (OSP) is a free, bilingual, web-based resource that provides teachers and students with clear descriptions of the “Essential Skills” and work habits important in work, learning, and life. Teachers planning programs in Canadian and world studies can engage students by using OSP tools and resources to show how what they learn in class can be applied in the workplace and in everyday life.

Skills identified in the OSP are:

A2.1 describe several ways in which geographic investigation can help them develop skills, including spatial technology skills and the essential skills in the Ontario Skills Passport (e.g., reading text, including graphic text; writing; oral communication; using maps, graphs, charts, tables; computer use; use of a geographic information system [GIS], satellite imagery; measurement and calculation; data analysis; decision making; planning; organizing; finding information; problem solving), that can be transferred to the world of work and to everyday life

Job Task Planning and Finding Information

D1.4 Data Analysis

identify the mode(s), if any, for various data sets presented in concrete graphs, pictographs, line plots, bar graphs, and tables, and explain what this measure indicates about the data

Tech (lots)

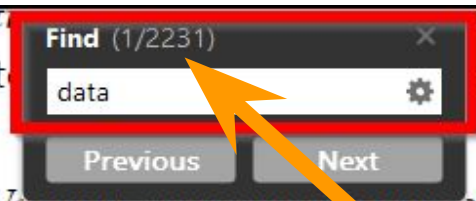
Ontario Skills Passport

A2.1 describe ways in which geographic investigation can help them develop skills, including spatial technology skills and the essential skills in the Ontario Skills Passport (e.g., reading text, including graphic text; writing; oral communication; using maps, graphs, charts, tables; computer use; use of geographic information systems [GIS], satellite imagery; measurement and calculation; data analysis; decision making; planning; organizing; finding information; problem solving), that can be transferred to the world of work and to everyday life

A2.1 describe ways in which geographic investigation can help them develop skills, including spatial technology skills and the essential skills in the Ontario Skills Passport (e.g., reading text, including graphic text; writing; oral communication; using maps, graphs, charts, and tables; computer use; use of a geographic information system [GIS], satellite imagery; measurement and calculation; data analysis; decision making; planning; organizing; finding information; problem solving), that can be transferred to the world of work and to everyday life

Search for “data” In All Courses (> 2k hits!)

- *As part of field studies:* In a field study, students can combine a number of ICT tools with field personal digital devices, and digital cameras.
- *As simulations:* Various simulation programs are available that provide hands-on visual engagement to support student learning.



Whenever appropriate, students should be encouraged to use ICT to support and communicate their learning. For example, students working individually or in groups can use computer technology and/or Internet websites to gain access to museums, galleries, archives, and heritage sites in Canada and around the world; databases of laws and regulations; political information and economic data on various countries; and digital atlases and other sources of information and data. They can also use portable storage devices to store information, as well as CD-ROM and DVD technologies, digital cameras, GIS maps, interactive whiteboards, and projectors to organize and present the results of their investigations to their classmates and others.

Grade 8 - Data Analysis with Code

Coding	<ul style="list-style-type: none">• No references to coding	<ul style="list-style-type: none">• Focus on coding throughout all grades, including:<ul style="list-style-type: none">◦ Gr. 1: Sequential events◦ Gr. 2: Sequential and concurrent events◦ Gr. 3: Sequential, concurrent, and repeating events◦ Gr. 4: Sequential, concurrent, repeating, and nested events◦ Gr. 5-6: Code involving conditional statements and other control structures◦ Gr. 7: Code involving events influenced by subprograms and other control structures◦ Gr. 8: Code involving the analysis of data
--------	---	--

First...

I am not an expert!

(but I'll do my best...)

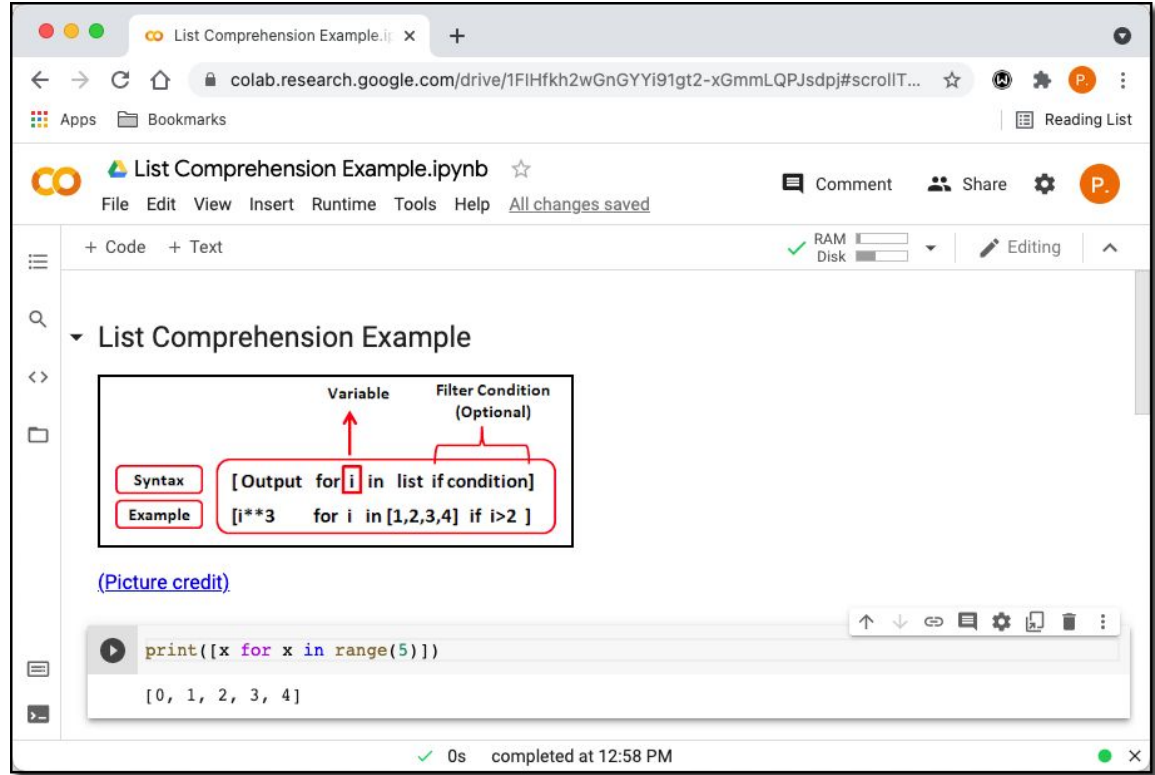
Second...

stackoverflow.com is your friend!

What is Colab?

Colaboratory, or “Colab” for short, is a product from Google Research.

Colab allows anybody to write and execute arbitrary **Python** code through the **browser**, and is especially well suited to machine learning, **data analysis** and **education**.



Why Colab is a Great Tool for Educators and Students

- By its very nature, Colab is a perfect tool for employing **Computational Thinking** techniques.
 - It is perfect for approaching problems in an iterative and step-by-step fashion, though its use of both text and code cells.
- It's from **Google**, so most students and educators already have an account.
- Integrated with **GitHub**
- It is an adaptation of **Jupyter Notebook**, which is very popular for scientific note-taking and interactive experimentation with code

Quick Demo of Colab

is.gd/BeensColabDemo

colab.research.google.com/notebooks/intro.ipynb#recent=true

Welcome To Colaboratory

File Edit View Insert Runtime Tools Help

Table of contents

- Getting started
- Data science
- Machine learning
- More Resources
- Machine Learning Examples

Section

What is Colaboratory?

Colaboratory, or "Colab" for short, allows you to write and execute Python in your browser, with

- Zero configuration required

Examples Recent Google Drive GitHub Upload

Filter notebooks

Title	Last opened	First opened	
Welcome To Colaboratory	8:28 PM	8:28 PM	

NEW NOTEBOOK CANCEL

604800

Colab notebooks allow you to combine **executable code** and **rich text** in a single document, along with **images**, **HTML**, **LaTeX** and more. When you create your own Colab notebooks, they are stored in your Google Drive account. You can easily share your Colab notebooks with co-workers or friends, allowing them to comment on your notebooks or even edit them. To learn more, see [Overview of Colab](#). To create a new Colab notebook you can use the File menu above, or use the following link: [create a new Colab notebook](#).

Colab notebooks are Jupyter notebooks that are hosted by Colab. To learn more about the Jupyter project, see

colab.research.google.com

IPO Model





- Input
 - Define the Objective
 - Get the Data
- Process
 - Have a Quick Look at the Data
 - Clean/Fix/Filter the Data (if necessary)
- Output
 - Plot and/or export the Data
 - Save chart picture

Relating IPO to Data Analysis: Input

The input come from outside the document:

- Google Sheets
- Excel
- Websites (a table)
- A file

...or inside the document

- From a list or a dictionary or other more complex data structures

Relating IPO to Data Analysis: Process

- Some examples:
 - Delete extra rows
 - Delete strange characters such as end- or footnote numbers
 - Multi-row headers
 - Take care of missing data (NaN = Not a Number)

Relating IPO to Data Analysis: Output

- Charts:
 - Bar (vertical and horizontal)
 - Scatter
 - Line
 - Histogram
 - Etc.
- Export to a new data file

Demos:

Working Through Some Real Examples

Get Data from Google Sheets

[Demo_Input_Data_from_Google_Sheet.ipynb](#)

This simple spreadsheet will let us plot x vs y data.

Years Taught vs Grade Taught (from Google Sheets)

[Demo_Teacher_Years_vs_Grade.ipynb](#)

This simple spreadsheet, collected from a Google Form will let us plot a scatter plot.

[[Colab Link](#)]

Get Data from Table on Website

[Demo_Input_Data_from_Table_on_Website_\(CAFE\).ipynb](#)

This is CAFE (Corporate Average Fuel Economy) data that needs some fixing before it can be graphed.

Adding a New Column using Math

[Demo Adding a New Column using Math.ipynb](#)

Doing some simple math to create a new column of data.

Bar Chart Demo (Mass of Planet)

[Demo_Bar_Chart_Mass_of_Planets.ipynb](#)

This data is taken from a table in Wikipedia and has to be cleaned up a bit before it can be graphed.

Tip!

I recommend reading in data from somewhere online.

When notebooks get moved or copied, their reference to any local data files could be lost.

GitHub is my favourite location for files but Google Sheets is equally easy to use as a data source.

Resource for Fixing Data

[Demo_Fixing_Data.ipynb](#)

A very helpful resource. Be sure to open it in Colab so you can see the Table of Contents.

[[Colab Link](#)]

The Template

Colab: is.gd/BeensDataTemplate

GitHub: bit.ly/beensdatatemplategithub

The GitHub Repository for Today's Resources

<https://bit.ly/3gZNPIN>

(github.com/pbeens/OTF-Data-Analysis-2021-05)

GitHub

(if time permits)

What is GitHub?

Some technical definitions (ignore!):

“GitHub is a **code hosting platform** for version control and **collaboration**. It lets you and others work together on projects from anywhere.”

--<https://guides.github.com/activities/hello-world/>

At a high level, GitHub is a **website and cloud-based service** that helps developers **store and manage their code**, as well as track and control changes to their code.

--<https://kinsta.com/knowledgebase/what-is-github/>

My definition:

A good place to share your data files and notebook (Colab) templates from.

github.com/pbeens/OTF-Data-Analysis-2021-05

Search or jump to... Pull requests Issues Marketplace Explore

pbeens / OTF-Data-Analysis-2021-05

Unwatch 1 Unstar 1 Fork 0

<> Code Issues 1 Pull requests Actions Projects Wiki Security Insights Settings

main 1 branch 0 tags Go to file Add file Code

pbeens Merge branch 'main' of https://github.com/pbeens/OTF-D... 6b304ef 1 hour ago 89 commits

datafiles	Create bad_cafe_data.csv	yesterday
.gitignore	Update .gitignore	13 days ago
Data_Analysis_Template.ipynb	Created using Colaboratory	1 hour ago
Demo_Bar_Chart_Mass_of_Pla...	Created using Colaboratory	1 hour ago
Demo_Fixing_Data.ipynb	Created using Colaboratory	17 hours ago
Demo_Input_Data_from_Goo...	Created using Colaboratory	22 hours ago
Demo_Input_Data_from_Table...	Created using Colaboratory	22 hours ago
README.md	Update README.md	1 hour ago

README.md

OTF-Data-Analysis-2021-05

This repository is for the OTF Data Analysis webinar on May 6, 2021, presented by Peter Beens (@pbeens) and Grant Hutchison (@grant_hutchison).

The presentation is a live demo of Google Colab, Google's implementation of Jupyter Notebook. We will learn how Colab is an excellent tool for the "Computational Thinking" process, in that it's an ideal way to analyze and document a problem and formulate a programming solution.

Table of Contents

- OTF-Data-Analysis-2021-05
 - Table of Contents
 - Basic Process
 - Data Sources

<https://bit.ly/3gZNPIN>

Current Repository
OTF-Data-Analysis-2021-05

Current Branch
main

Push origin
Last fetched 4 minutes a... 1 ↑

Changes

History

0 changed files

No local changes

There are no uncommitted changes in this repository. Here are some friendly suggestions for what to do next.

Push commits to the origin remote

You have 1 local commit waiting to be pushed to GitHub.

Always available in the toolbar when there are local commits waiting to be pushed or P

Push origin

Open the repository in your external editor

Select your editor in [Preferences](#)

Repository menu or A

Open in Visual Studio Code

View the files of your repository in Finder

Repository menu or F

Show in Finder

Open the repository page on GitHub in your browser

Repository menu or G

View on GitHub

Summary (required)

Description

Commit to main

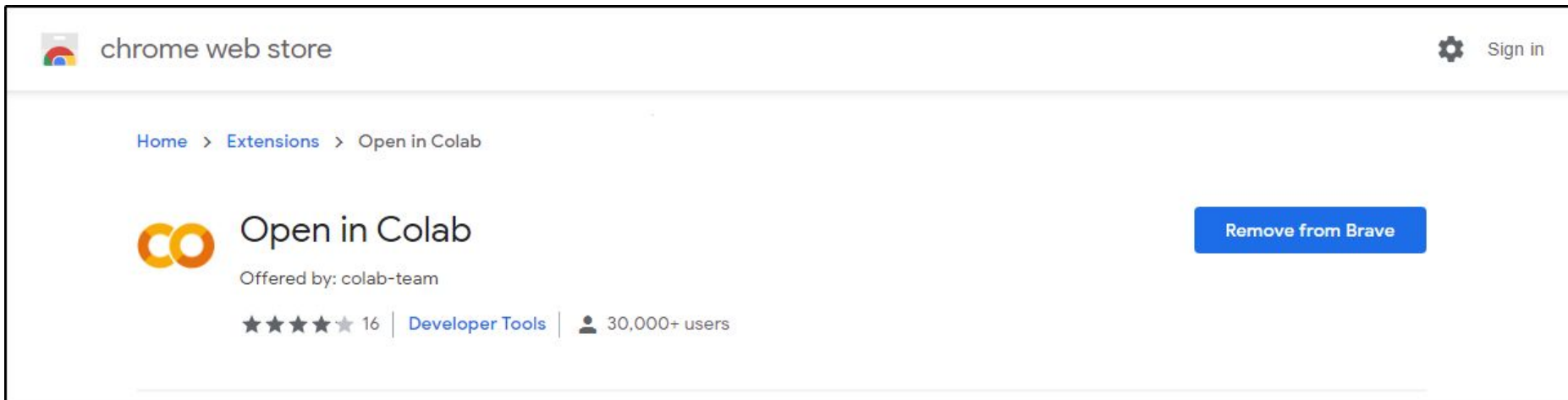
Committed just now
Updated the basic process

Undo

GitHub Desktop - <https://desktop.github.com>

Open in Colab Chrome Extension

Allows you to open any Jupyter or Colab notebooks in GitHub directly in Colab.



Callysto ([Callysto.ca](https://callysto.ca))

- Is an alternative to Colab
- Many resources across subject areas
- Has many training materials
- Hosted in Canada
- Uses Jupyter notebooks in a similar fashion to Colab, but is not integrated with Google or GitHub
- Being used for the upcoming [Hackathon for Teachers](#) by Andrew McDonald ([@afmcdnL](#))

ACSE Mail List:

<https://acse.net/mail-list/>



Questions?