

Data Basic.io

Data Science

Cognitive Constructivism

[DataBasic.io](#) is a suite of easy-to-use web tools for beginners to explore different sets of data. It is a project from the Engagement Lab at Emerson College and the MIT Center for Civic Media.

DataBasic features multiple tools to help you understand textual and tabular (tables) data, including WordCounter, samediff, WTFcsv, and ConnectTheDots. The tools were tested in classrooms and workshops to make sure they worked well and were easily understandable. Users can visualize the resulted information in a variety of useful formats that facilitate better understanding of the concepts (e.g., color-differentiated tables with graphic cues that can increase retention).

- [Data Basic Word Counter](#): Word Counter tells you the most used words and phrases in a document. It helps you quantitatively summarize your text so you can investigate how words are used. It creates a basic word cloud, but also does word counts, in several new interesting ways by breaking its analysis into a few different pieces – A word cloud, Top Words, Bigrams, and Trigrams.
- [WTF CSV](#): WTF CSV gives you a quick overview of a CSV file you upload. It's designed for the data newbie who has no idea what to do with a spreadsheet. It helps a user peek behind all those columns and rows through some simple analysis by giving you the frequencies of each attribute (column) in a table, presenting it as a histogram or bar chart or word cloud depending which visualisation is most appropriate. It also provides a "What do I do next?" set of questions that can help prod the beginner.
- [Same Diff](#): SameDiff lets you compare two documents to see how they are similar and different. It helps you see differences and similarities in the words used in each file so you can learn about quantitative analysis of text. SameDiff compares one corpus (a collection of written text) of text to another corpus of text to show you similarities and differences. It uses a cosine similarity algorithm (based on the number of times words are used in each) to rate whether the documents are really similar or totally different.
- [Connect the Dots](#): Connect the Dots shows you how your data is connected by analyzing it as a network.



[Watch on Vimeo](#)

Tool Snapshot

| | |
|------------------------|---|
| Price | Free |
| Learning | Cognitive Constructivism |
| Ease of Use | ★★★★☆ |
| Privacy | ★★★★★ |
| Accessibility | ★★★★★ |
| Class Size | Unlimited |
| ISTE*S | Knowledge Constructor & Computational Thinker |

Type of learning: Students learn to analyze and construct meaning from data.

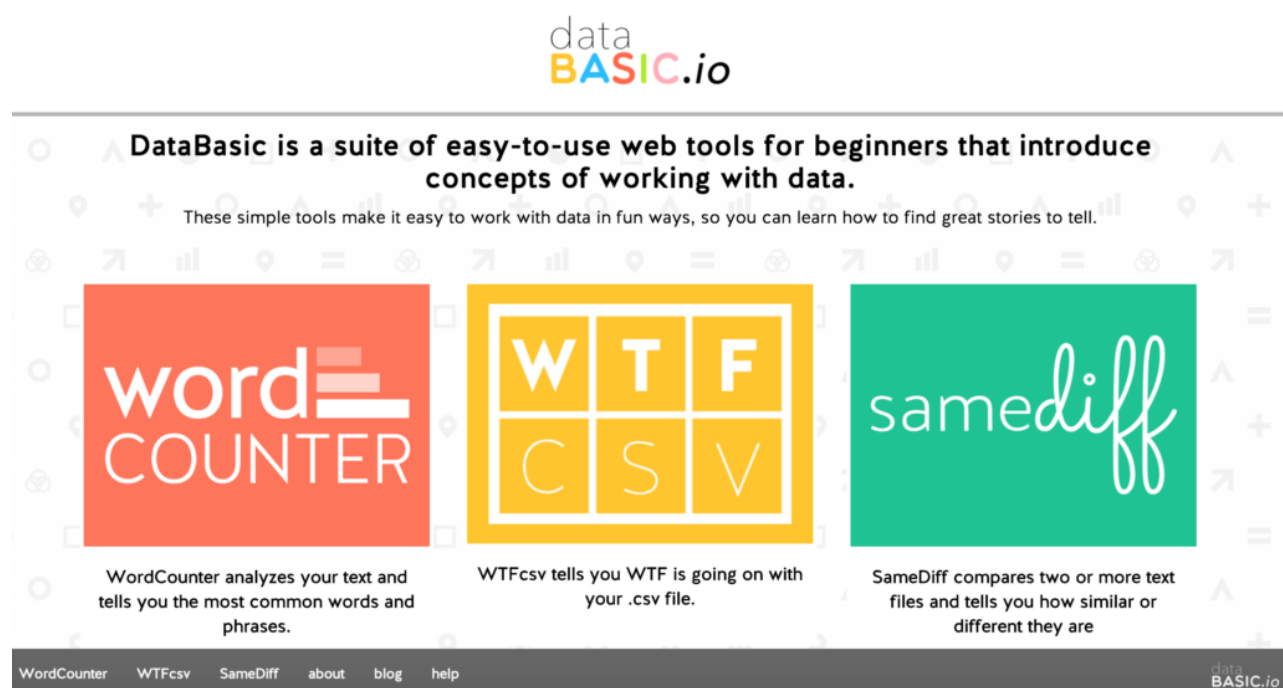
Privacy: The Data Basic website stores information uploaded for only the amount of time it takes users to analyze the data, then it is deleted. The aggregate results shown (metadata) are kept for 60 days, and then get deleted.

Ease of Use: Tool has simple and easy-to-use interface.

Accessibility: Data Basic supports screen readers.

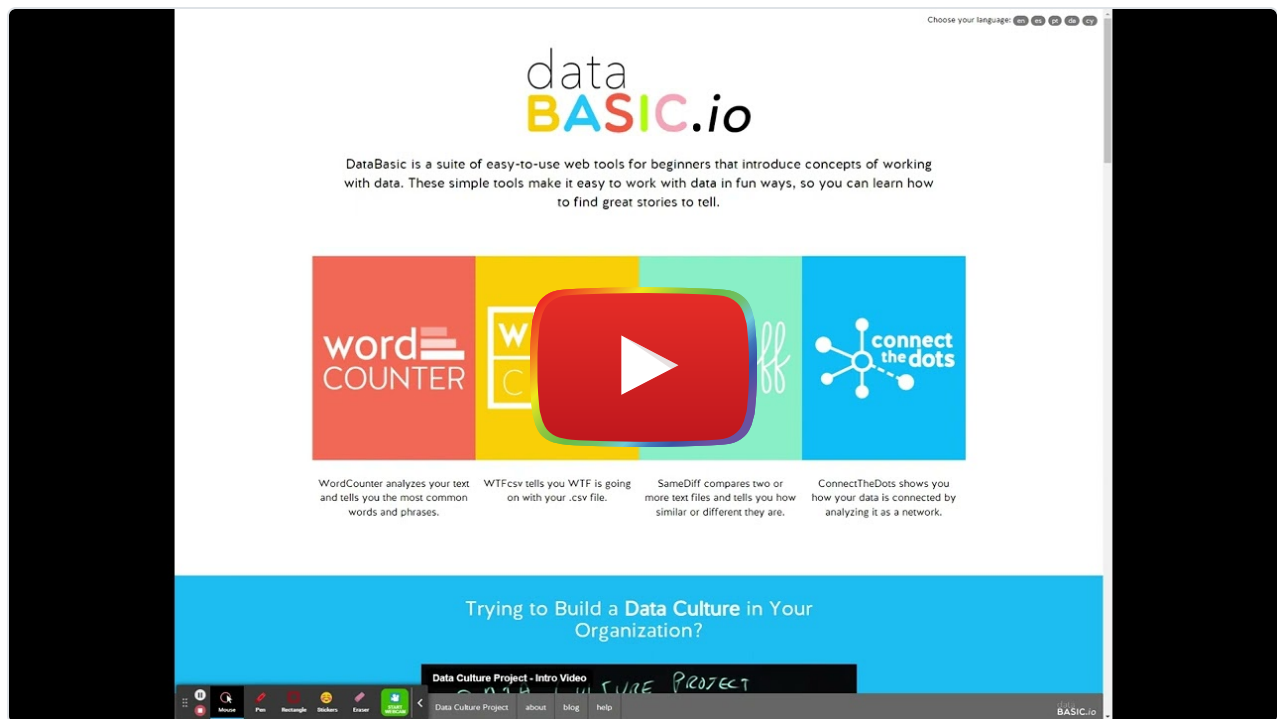
Power and Bias: Website is available for English, Spanish, and Portuguese speakers.

[ISTE Standards for Students](#): Hands-on approaches to analyzing data. Students construct their knowledge through collecting data and trying to find stories and information within it.



Main page of the website databasic.io

DataBasic.io Overview Video



[Watch on YouTube](#)

Data Basic & the SAMR Model

Dr. Ruben Puentedura's [SAMR model](#) offers a lens for examining how technology is adopted in a classroom. As you strive to incorporate online tools into your classroom, we encourage you to use this model as an analytic tool.

Here is an example of how Data Basic might fit within the SAMR model:

- **Substitution:** Students upload a dataset to WTFcsv to get a series of visual charts rather than designing their own charts on paper in class.
- **Augmentation:** Word Counter instantly generates word clouds and bi-/trigrams of most popular words in a text, which provides students with a new way to analyze textual data.
- **Modification:** Using Samediff, students can instantly compare and contrast two different text documents.
- **Redefinition:** Students can make hypotheses about textual data and get immediate results to confirm or reject their hypotheses.

Far too often, technology is used as a direct substitute for other low-tech tools (e.g., pencil and paper). While substitution has some benefits (e.g., students develop their technology skills and knowledge), we encourage you to think about how you might use Twitter to modify or redefine learning.

Learning Activities

[WordCounter Activity Guide](#): Use **WordCounter** to analyze a musician's lyrics and sketch a story with their words. Sketching a data story helps participants build their data literacy by finding and drawing a story about what words show up in famous musicians' lyrics.

[WTFcsv Activity Guide](#): Use **WTFcsv** to identify some questions to ask your spreadsheet. This hands-on activity helps participants build their data literacy while looking at some fun datasets and brainstorming questions you can ask about them.

[SameDiff Activity Guide](#): Write Your Own Song Use **SameDiff** to analyze the lyrical styles of two musicians and invent a duet for them.

Lesson Ideas Using WordCounter

Personal Narratives

Write, or copy and paste, a personal narrative into WordCounter. Students will be able to see common themes from the visuals that are produced from the Data Basic WordCounter's word clouds, Top Words, Bigrams, and Trigrams. These word clouds make great displays for a classroom bulletin board.

Current Affairs Analysis

Copy and paste a news story into Data Basic WordCounter. Ask students to guess what the story was about and generate a headline based on the results.

Student Profiles

Have all students in your class write a few positive adjectives about each of their classmates anonymously. Compile all the papers, input the adjectives for each student into WordCounter, and generate a student profile word cloud to give back to the student.

Collaborative Group Work

For the solar system, student teams could each be responsible for creating Word Clouds to describe one of the eight planets or in a Math class student teams could be responsible for creating a visual representation listing all the types of

quadrilaterals.

Lesson Ideas Using SameDiff

Compare and Contrast: Compare and contrast the documents of two or more students' writing, famous speeches, song lyrics, news reports, or book reviews.

How to Use DataBasic

- Go to databasic.io
- Click one of the three tools: Word Counter, WTFcsv or SameDiff
- Word Counter: You can use any sample from drop down menu, paste a url, paste a text or upload your own file. Hit Count button to get started .
- WTFcsv : You can use an existing sample file from the drop down menu or upload your own file. Hit Analyze button to continue exploring.
- **SameDiff** : You can use an existing sample file from the drop down menu or upload your own files/speeches. Hit **Compare** button to continue exploring. You can go back and forth between each of data basic tool by clicking on the menu at the bottom.

Research

D'Ignazio, C., & Bhargava, R. (2016). [DataBasic: Design principles, tools and activities for data literacy learners](#). *The Journal of Community Informatics*, 12(3).

D'Ignazio, C. (2017). [Creative data literacy: Bridging the gap between the data-haves and data-have nots](#). *Information Design Journal*, 23(1), 6-18.



This content is provided to you freely by EdTech Books.

Access it online or download it at <https://edtechbooks.org/onlinetools/data-basic-io>.

