How to Clean the Data

1. Make sure Python 3.7 is installed on the computer (https://www.python.org/downloads/)
2. Install Anaconda and Spyder (You can use juPyter if you want) https://www.anaconda.com/distribution/ and https://docs.spyder-ide.org/installation.html
3. Open TechScript.py
4. Have the survey results formatted correctly (see below for specifics: **How to Format the Data**) and saved as a comma separated value (.csv) file
5. Ensure that TechScript.py and the .csv file are in the same folder
6. Change the pink text in TechScript.py to whatever the .csv file with the data is named
7. Change the red text in TechScript.py to the directory you want to save the cleaned data to

How to Format the Data

1. In an Excel document, name 8 columns “ID”, “Relationship”, “Meaning”, “Topics”, “Initiatives”, “Session Preference”, “Resources”, and “Focus Group”. These will correspond to the questions asked in the survey.
2. In the “ID” category, input the survey response number as shown in the form
3. Copy the responses for all the other categories, up to “Focus Group”
4. For “Focus Group”, input a 1 if the individual wants to be part of a Focus Group and a 0 otherwise
5. Remove any redundant data (people who answered twice), if applicable
6. A finished row will look like this

	A	B	C	D	E	F	G	H
1	ID	Relationship	Meaning	Topics	Initiatives	Session Preference	Resources	Focus Group
2	1130	DTX Launch Detroit Student Accelerator Participant D-Venture Entrepreneur-in-residence (EIR) program Participant DTX Fellows Program Participant	Learning about new TechTown initiatives Hearing about upcoming TechTown programs Receiving information before anyone else Being part of a group	Resources and opportunities from TechTown partners Neighborhood news (community updates, grand openings, anniversaries, events, etc.) Tips and tricks from TechTown staff and other industry leaders Detroit Events (Toast of the Town, Rock Your Guac, etc.)	Attend alumni only workshops Access to alumni only office hours Utilize 1-1 support (mentoring / coaching from TechTown or others in our network)	Either	Regular TechTown Newsletter Flyers + Other Information at TechTown	1

How to Generate the Probability Matrix

Background: What does this show?

Each graph will compare the likelihoods of question responses with regard to what tech program survey respondents participated in.

1. To find the relationship between the program and a question, put a # in front of the desired paragraph (it will be one of the paragraphs in blue). Each paragraph corresponds to dropping a survey question from the matrix (the question topic is listed above each paragraph). Each phrase in green represents an option you can choose within the given question.

The section with a # in front of it will show up on the matrix. This is because the # in front of the line prevents that section from being dropped (“#corr.drop”). Everything else (without a #) will be removed.

2. It will look like this:

```
106 #corr.drop(['Incubation Center', 'DTX Launch', 'EIR Program', 'DTX Fellow
107
108 corr.drop(['New Initiatives', 'Upcoming Programs', 'First to Get Info', '
109
110 corr.drop(['TechTown Resources', 'Partner Resources', 'Neighborhood News'
111
112 corr.drop(['Attend Event', 'Attend Workshop', 'Alumni Workshop', 'Alumni
```

3. Copy and paste TechScript.py into Spyder

4. Press the green triangle. Voila.