**CRNY CONTEST SUBMISSION**

|  |  |  |
| --- | --- | --- |
| **Name of Teammate** | Gauri Chaudhari | Radhika Agarwal |
| **Contact Information** | [gchaudh@iu.edu](mailto:gchaudh@iu.edu) | [radagar@iu.edu](mailto:radagar@iu.edu) |
| 812-803-9913 | 812-803-8575 |
| **College & Department** | Luddy – Data Science | Luddy – Data Science |
| **Student Status** | Graduate | Graduate |

**Title of Work:** A Visual Impact on New York’s Creative Pulse

**Link to Data Visualization:** <https://public.tableau.com/app/profile/gauri.chaudhari/viz/AVisualImpactonNewYorksCreativePulse/CRNY>

**A 100–word narrative describing what the data visualization is about:**

"A Visual Impact on New York’s Creative Pulse" encapsulates the essence of the city's artistic reawakening, depicted through an innovative interactive dashboard fueled by the Guaranteed Income for Artists program. This visualization not only highlights the demographic diversity and narratives of New York's artists but also brings to the forefront the profound interplay of identity, inclusivity, and artistic endeavors. At its core, the map visualization stands as a testament to this revival, offering users a dynamic and intuitive exploration of the cultural landscape. It magnificently charts the geographical distribution of artists, their engagement levels, and the vibrant mosaic of their collective experiences, making it a pivotal element in understanding New York's creative resurgence.

**Description of the methods used, including any data formatting/recoding:**

In crafting our interactive tableau data visualization, we relied on the Guaranteed Income dataset as our foundation. Our methodology was meticulous, prioritizing data integrity through extensive cleaning(especially on the geographical aspects) and preparation processes.

The approaches performed in Python include:

* The completely blank rows except the ID are essentially the participants who started the application but did not fill in any other details. We assumed they were invalid entries and removed those rows leaving us with 21921 participants out of 23797 participants in the total dataset.
* Converted the datatype of zipcode to an integer since Tableau uses zipcode to outline latitude and longitude
* Removed whitespaces and trimmed all cells of the data
* Modified the case of the city to match TitleCase

The approaches performed in Excel include:

Our strategy in Excel involved disaggregating multi-response variables across separate sheets, linked via participant IDs. This crucial step shifted our data from a wide to a long format, maintaining the integrity of each participant's responses across multiple variables. Geographic data underwent a rigorous validation process, cross-referencing zip codes with a detailed lookup table (<https://simplemaps.com/data/us-counties>) to correct city names and ensure geographical accuracy.

This also solved the issues of incorrect spellings in the cities entered by the participants.

The approaches performed in Tableau for data preparation include:

* For some cities, tableau was displaying it as unknown so as an alternative to get complete data we googled the latitude and the longitudes and updated it within Tableau since Tableau internally converts all zipcodes to latitudes and longitudes (<https://simplemaps.com/data/us-counties>)

For our final visualization, we have used tableau –

1. We used calculated fields to calculate metrics like total valid applicants, enrolled participants, not enrolled participants, and withdrawn participants as well as for the overall artist individuality metrics like percent of immigrants, LGBTQIAP+, specially-abled, or participants with a criminal legal history.
2. Our map visualizations leveraged Tableau's dynamic zone visibility, enabling users to drill down from regions to counties, and counties to cities with seamless interactivity. (<https://help.tableau.com/current/online/en-us/dynamic_zone_visibility.htm>)
   1. We made use of parameters and actions in Tableau to create interactivity between region, county, and city maps
   2. This gave us a unique ability to click on a particular region so that we can view its respective counties and by clicking on counties we can view the cities of those counties. Simultaneously we can view the demographic information as well as the count of enrolled participants for the selected geographical area by hovering.
   3. For the hovering effect, we used viz in the tooltip- (<https://help.tableau.com/current/pro/desktop/en-us/viz_in_tooltip.htm> )
   4. The end user can select from the demographic dropdown (which are our filters provided at the left of the map).
3. To explore the overall New York City nature of the applicants and for the variables that were multi-response we created a narrative of visualizations such as bubble charts, treemaps, etc

**List of any tools/software used:**

* Tableau Public
* Python 3.11.4
* Microsoft Excel
* Visual Studio Code

**If applicable, a copy of the code used to create the visualization:**

<https://indiana-my.sharepoint.com/:u:/g/personal/gchaudh_iu_edu/ETPkbNH8Kq9Iub7ZECVKmQMBbkdNo2lPOdkk_dKO_Io6bA?e=h4sR4t>