Exploring the Effectiveness of Icons/Emoji in Data Visualization and Data Communication: Evaluation Plan

Zach Palmer & Anmol Paudel

Introduction

This project is an exploration in determining whether icons can successfully communicate relevant information through data visualizations. The purpose of this experiment is to determine the feasibility of using emoji or icons to replace traditional legends on maps.

We intend to explore the following questions:

1. Can icons/emoji communicate the same information as maps with legends?
2. Are people comfortable/confident with interpreting maps that use icons/emoji?
3. Are there differences in time spent looking at maps that make use of emoji vs. Legends?

Methodology

We will construct a JavaScript web app that contains multiple maps for participants to browse and study. The maps will be populated with data from two datasets containing data collected in San Francisco. They include data on housing development, and police incident reports. This experiment will use MTurk to recruit participants. Participants will then be split into the two following groups for A/B testing:

Group A – Maps will display data using more standard heatmaps and legends

Group B – Maps will display data using icons/emoji instead of legends

MTurk participants will be asked to study the maps and then answer questions about the maps and their confidence in their answers to gauge how well they were able to interpret the maps, as well as how confident they are in their answers. The questions will each contain two parts:

1. The first half will be a simple multiple-choice question
    i. Ex: What quadrant of the map appears to be the noisiest?
2. The second half will use a 5-point Likert scale to ask participants to rank how confident they are in their answers.
    i. Very unconfident, somewhat unconfident, neutral, somewhat confident, very confident

Their answers will be collected in order to assess participant accuracy in interpreting the maps, as well as their confidence in their answers.

Hypotheses

The following are the hypotheses for the three research questions posed in the introduction of this document:

First Hypothesis:

H_0: There is no difference in scores[1] between Group A and Group B.

H_a: There exists a difference in scores between Group A and Group B.

Second Hypothesis:

H_0: There is no difference in confidence[2] between Group A and Group B.

H_a: There exists a difference in confidence between Group A and Group B.

Third Hypothesis:

H_0: There is no difference in time spent viewing maps between Group A and Group B.

H_a: There exists a difference in time spent on maps between Group A and Group B.

Activity

1. Participant will accept the job on MTurk.
2. Participant will be redirected to Qualtrics.
3. Participant will sign the informed consent form.
4. Participant will fill out basic demographic information.
5. Participant will do the 2 questionnaires. One with Type A map and other with Type B map. The maps will be embedded as iframes within the questionnaires.
6. A completion code will be generated when the participant completes the survey on Qualtrics which will be needed by MTurk.

Questionnaire

The demographic questionnaire will ask about

Age (select from groups)

- 18-25
- 25-35
- 35-50
- 50+

Gender (select from groups)

- Male
- Female
- Other
- Rather not disclose

Education Level

- Less than High School
- High School Completed
- Some college or currently enrolled

[1] Here score refers to the number of questions answered correctly (Normalized)

2 Here confidence refers to average confidence reported in responses (Normalized)

- Bachelor's Degree Completed

We will have 4 Questionnaires in the survey, 2 surveys with traditional maps and 2 surveys with our experimental maps. Each survey will have 4 questions with a multiple-choice response and a corresponding confidence scale.

The questions that will be asked about the two datasets are as follows:

- **Housing development**
  - Which areas are more developing?
  - Which areas are less developing?
  - How developed would you say area X is compared to area Y?
  - Can you guess, in number of units completed, how many of them were in Area Z?
    - (Options in scales like 10-50, 100-500, 1000-5000)
- **Incident Reports (Crimes)**
  - Which areas have more incidents?
  - Which areas have less incidents?
  - How safe would you say area X is compared to area Y?
  - Can you guess, in incidents per section how many incidents were reported in Area Z?
    - (Options in scales like 10-50, 100-500, 1000-5000)

Each question will be associated with a confidence slider scale.

How confident are you of this answer? How many of the choices that you thought were wrong and eliminated them?

0. Very unconfident – I randomly choose one answer, I couldn't eliminate a single choice

1. Unconfident – It was a choice among 3 options, I could just eliminate one of the choices

2. Neutral - It was a choice among 2, I could eliminate at least two of the choices.

3. Confident - I am absolutely sure of my choice. I could eliminate all other choices.

Timing data will be collected for different questions using in-built features from Qualtrics.

Data Collection

Data we will be collecting:

- Age, location, gender, educational background (if provided by MTurk system)
- Time spent viewing the maps
- Questions that have been answered incorrectly
- Questions that have been answered correctly
- For incorrectly answered questions, the incorrect choice
- Confidence in each response

We hope to collect about 100 responses split into 2 groups. Group 1 will see housing dataset on Type A map and Incident dataset on Type B map. Group 2 will see housing dataset on Type B map and Incident dataset on Type A map.

Analysis and Interpretations

First, we will plot/visualize the data to see what kinds of pattern emerges and what kind of statistical tests can be done with the data collected.

But for now, just to pre-register, our plan is to do

1. Two tailed independent sample t-test for the first hypothesis
    i. Difference of means between group G1 and group G2 for the N participants
2. Two tailed paired sample t-test for the second hypothesis
    i. Difference of confidence between Type A maps and Type B maps, but paired for the N participants
3. 2-way Anova for the third hypothesis
    i. With difference in timing intervals between answering each question
    ii. Ti will give the time interval between answering Question Qi and Q(i+1)
    iii. Post-hoc analysis depending on the results that we'll get