

Experimental_Methodology

Mark Paluta, Krysten Thompson, Chris Ventura

March 10, 2019

Research Question

How does age (and maybe other factors) affect the chance of receiving a reply when reaching out regarding a roommate posting on Craigslist?

Treatment Conditions

- Age (what ages do we want to use?)
- any others?

Methodology

The pilot study will last 7 (?) days. Each day, once per evening, we will count the number of posts from that calendar day. We will randomly assign an array of that length to the number of treatment conditions. For each posting, we will note several covariates and we will give up to 48 (?) hours for a response. All postings will be considered valid except for:

- Posts specifically seeking a certain age or sex (?)
- Duplicate postings as recognized via images, post title, or body text

Data Collection

We will collect a few types of information:

- 1) Metadata about the post for purposes of avoiding double-counting and archiving in case the post is taken down.

```
colnames(df)[1:3]
```

```
## [1] "Posting.URL"      "Post.Title"      "Posting.Body.Text"
```

- 2) Features of the post that may itself predict high or low response rate

```
colnames(df)[4:7]
```

```
## [1] "Posting.age"      "Listing.Price"    "Listing.Type"     "Listing.Market"
```

- 3) Data about the poster, which may also influence response rates to treatment

```
colnames(df)[8:9]
```

```
## [1] "Poster.Gender"    "Poster.Age"
```

- 4) Our treatment(s)

```
colnames(df)[10:11]
```

```
## [1] "Treatment.Age"    "Treatment.Gender"
```

- 5) Any relevant metadata about our reply, which too could influence response rate

```
colnames(df)[12]
```

```
## [1] "Response.Time"
```

6) Our outcome variable

```
colnames(df)[13]
```

```
## [1] "Reply"
```