Capstone 2 Report

"Projecting Product-Customer Fit with Public Data"

Sam Fisher

Disclaimer: This project uses public data from Pew Research Center's "Information Engaged and Information Wary" survey.

Pew Research Center bears no responsibility for interpretations presented or conclusions reached based on analysis of the data.

Introduction

Our client, Edugames.org, is a fictional but plausible company that is interested in creating online math games for seniors. They have traditionally created sudoku puzzles and other paper media but are interested in creating a new offering for seniors in the online space. Their hope is that this new product will offer an enjoyable avenue for intellectual engagement and growth during one's professional retirement.

The current thinking is that the game will be a mobile and web app. The company wants to know if there are particular demographics amongst seniors that are more likely to become engaged users if they were advertised to and converted. *Since the product has not yet launched, we don't yet have hard evidence correlating demographics to conversions and engagement.*

Instead, we'll use public survey data about technology and information engagement from the Pew Research Center to measure behavioral traits that we assume will correlate to successful customer conversion and retention.

The Ideal Customer

The objective of this analysis is to identify demographic traits that correspond with customer fit. Let's begin with qualitative assumptions about who the ideal customer is.

We'll assume that the ideal customer is a senior (Age of 65+) who:

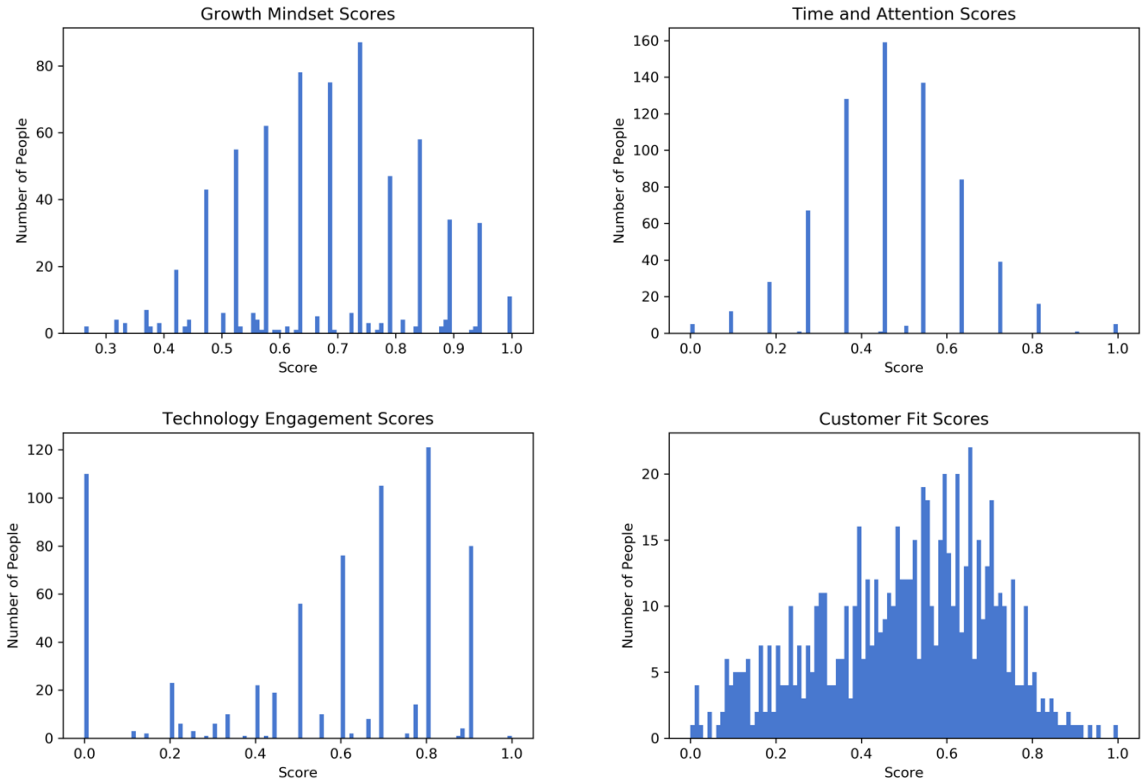
- Exemplifies the belief that their learning pursuit will be fruitful and that a new experience can be rewarding. [Growth Mindset]
- Has the time and attention necessary to play an involved and challenging game, potentially for multiple hours per day. [Time and Attention]
- Engages frequently with the internet and technology and has access to the necessary hardware to play the game. [Technology Engagement]

Quantifying Customer Fit Scores

The customer fit score will be our “target variable” in a regression analysis. We will try to predict this score for each senior given some demographic information about them. The coefficients of the linear model we fit to the data will tell us how each demographic feature relates to the customer fit score.

We begin by taking a sum of survey responses related to each of growth mindset, time and attention, and technology engagement. We then aggregated these individual features in order to create the general customer fit score.

Below you can see the distributions of these scores for subset of seniors whose responses were used in the model. Note that for each feature 0 is the least ideal customer and 1 is the most ideal customer.

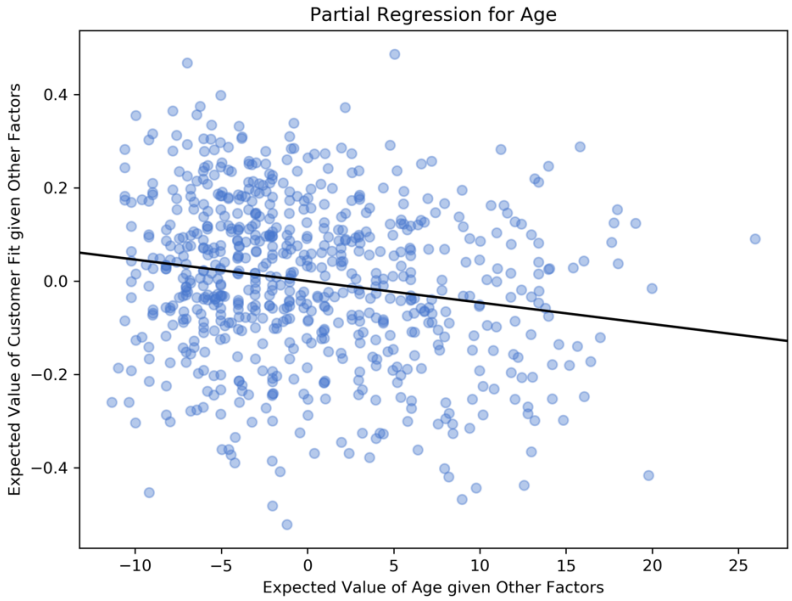


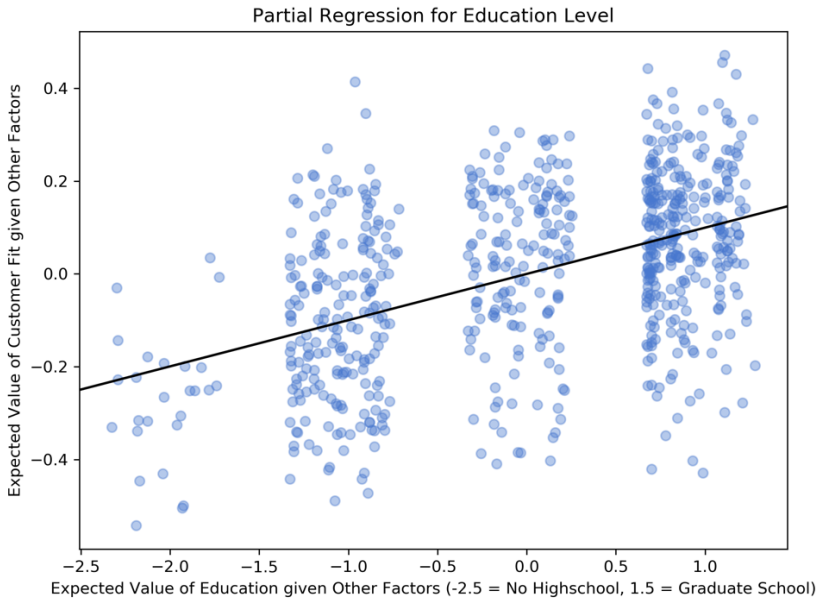
Notice how each component feature is visually expressed in the customer fit score. For the customer fit scores depicted (and used in the model) the scale has been adjusted such that the minimum score in this subset of individuals is 0 and the maximum is 1.

For more information on this process, see the “Creating Target Customer-Fit Features” section of the methodology notebook.

Recommendations to Edugames.org

We find that younger (ages 65-70), college-educated seniors are substantially more likely to fit our profile for the ideal customer. We recommend beginning marketing efforts with a focus on this subsection of the population. The following graphs depict the influence of age and education on customer fit in a linear regression model:

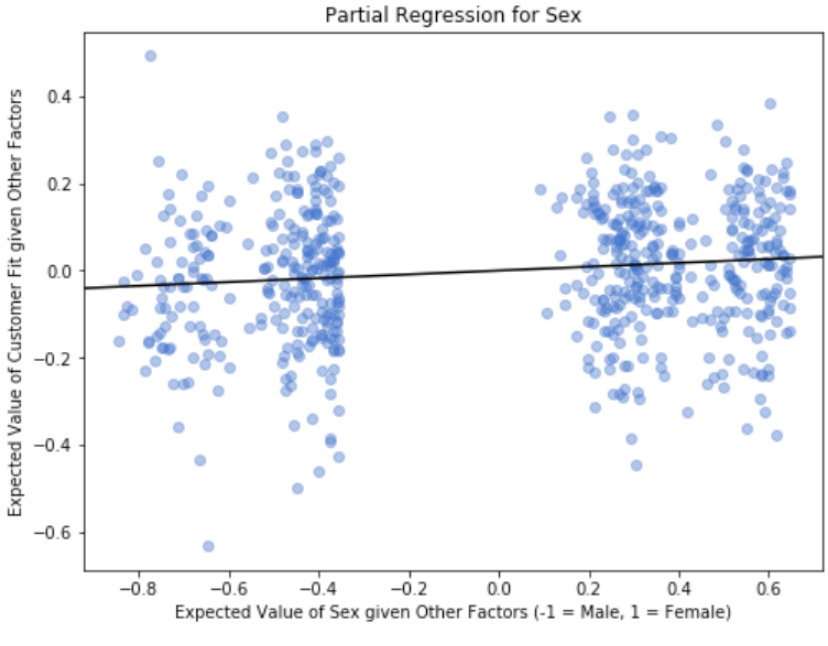


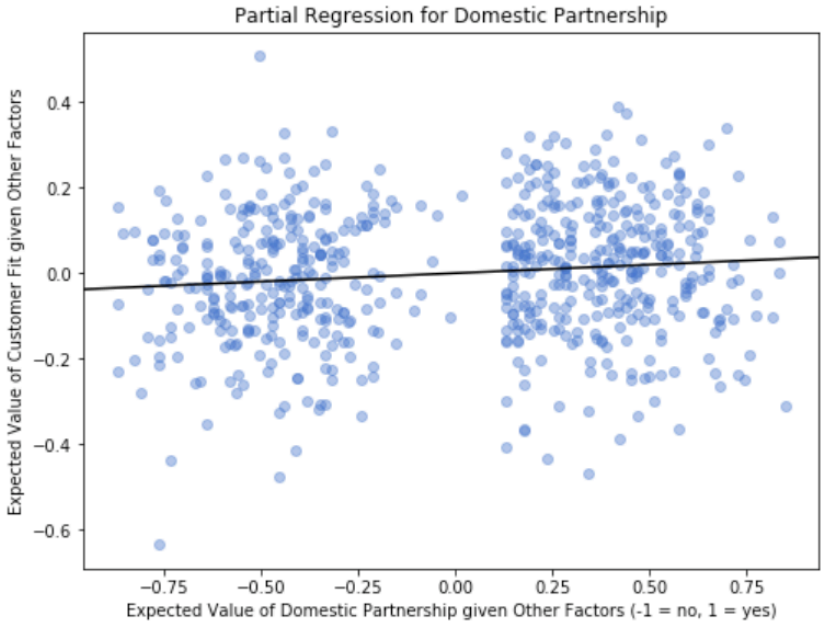


Additionally, we recommend implementing a/b tests for effects of the following variables on customer conversion, controlling for age and education level:

- 1) “Sex”: Customer is male or female
- 2) “Domestic Partnership”: Customer has a romantic partner that lives with them.

In our analysis, being female and being in a domestic partnership both increased the expected fit to our ideal customer profile in a small but statistically significant way. For domestic partnerships, this may simply be a correlate to an effect driven primarily by level of wealth.





Additional Observations from Modeling

Greater income correlates positively with being in a domestic partnership (spearman's rho value of 0.38) and education level (rho of 0.53). Distributing marketing efforts between wealthy and less wealthy people is a serious mission-related consideration in relying on education level or domestic partnerships to predict for customer fit. In order to mitigate the impact of targeting marketing towards wealthy people, Edugames.org may wish to pursue "outreach" focused marketing in parallel (likely with different strategies in mind.)

We found that the demographic indicators mentioned above were fairly predictive of technology engagement, explaining 36% of the variance, but did very little to account for the broad spectrum of survey responses related to growth mindset and time/attention, explaining only 8% and 2.5% of the variance respectively. A logistic regression model was able to classify with 72% accuracy whether a senior was a mobile internet user, and it's weighting of the covariates used echos that of our linear model above.

We expect that using online marketing channels would naturally introduce a helpful selection bias in the technology engagement of potential customers reached.

For more detail on how modeling was carried out, see the "Regression Analysis for Customer Fit" section of the methodology notebook. For more detail on the classification task, see "Classification for Insight on Mobile Use."

Exploratory Analysis for General Insight

Certain insights from this dataset can help us to frame a few expectations for the senior population as whole. These insights are less focused on promoting a specific action today and more focused on framing how we think about the landscape in which we're operating.

A) Technology Adoption

Let's begin by distilling relevant facts from Pew's Report on Senior Tech Adoption. See the linked report for greater detail on the insights below.

According to Pew's Report,

Seniors make up a substantial and growing portion of the population.

- Seniors (65+) make up 15% of the population- about 46 million people.
- The share of seniors is projected to increase to 22% by 2025.

Tech adoption for seniors is growing in step with the general population, especially for younger, wealthier, and college educated seniors.

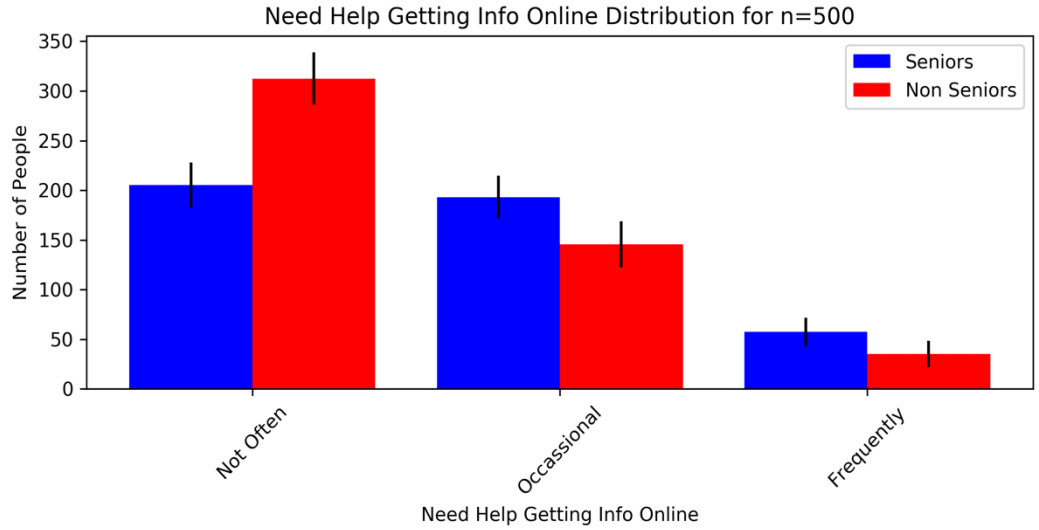
- Growth in internet, broadband, smartphone, tablet, and social media adoption amongst seniors closely mirrors the high rate of growth in the general population.
- The total rate of adoption of these technologies varies is a little less than 2/3 that of the general population.
- Younger, wealthier, and college educated seniors show higher rates of tech adoption in terms of internet, broadband, smartphone, tablet, and social media adoption.
- 40% of seniors own smart phones
- 25% of seniors play online games.

Seniors who do adopt internet and device usage have high levels of engagement.

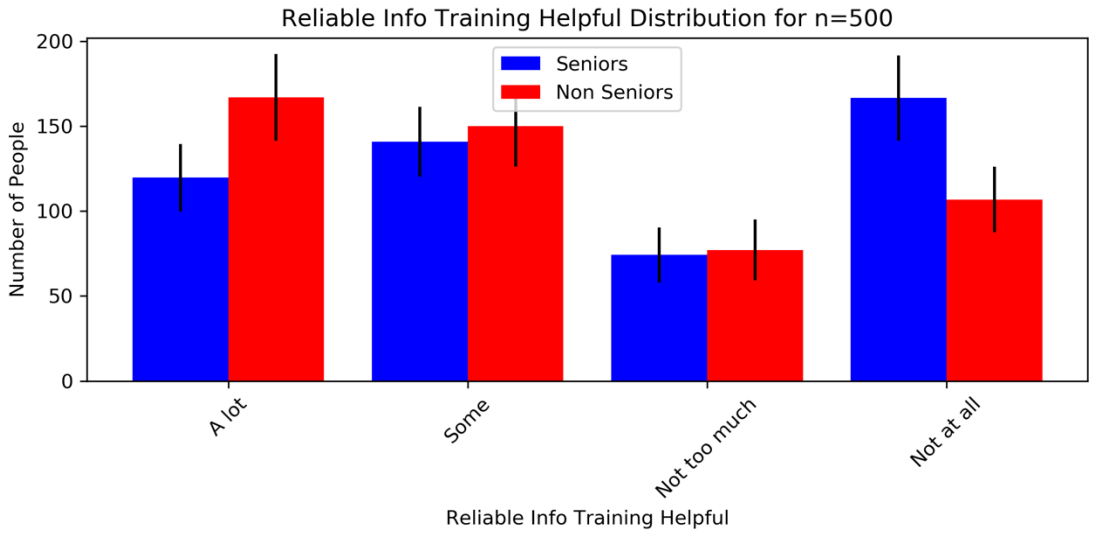
- 75% of internet-using seniors go online daily.
- 76% of smart-phone owning seniors use the internet multiple times per day.

B) Information Constraints & Confidence

About 3 of 5 seniors (age 65+) report needing occasional or frequent help getting information online, compared with about 2 of 5 people in the general population.

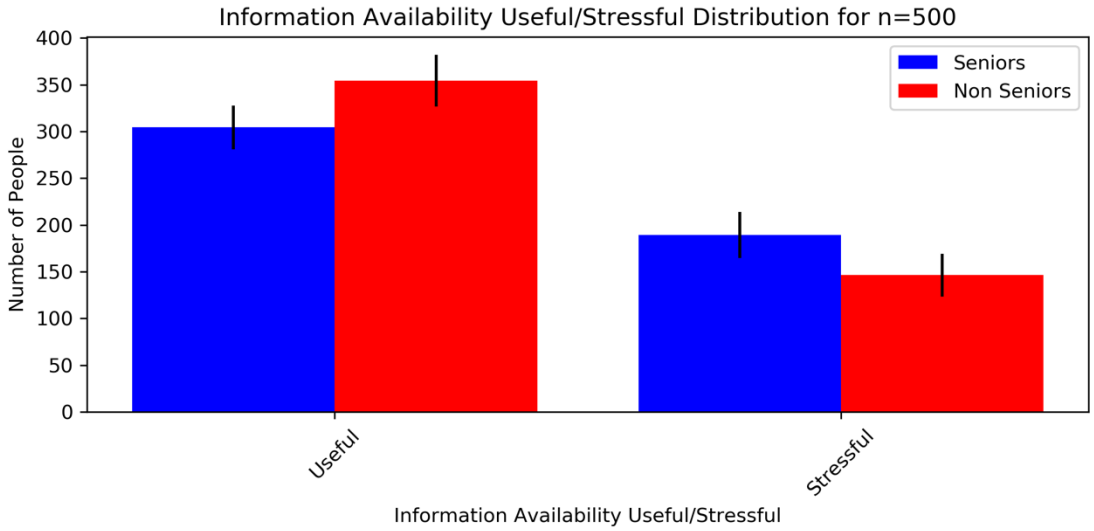


Conversely, seniors were fairly evenly divided on whether training on identifying reliable information would be helpful, perhaps demonstrating that some seniors are comfortable deciding which sources to believe.



From this we can infer that there are significant differences in the distributions of individuals who exemplify “information” and “technology” comfort. Information comfort is the degree to which one is able to confidently decide upon trustworthy sources. Technological comfort is the degree to which one is confident in using devices.

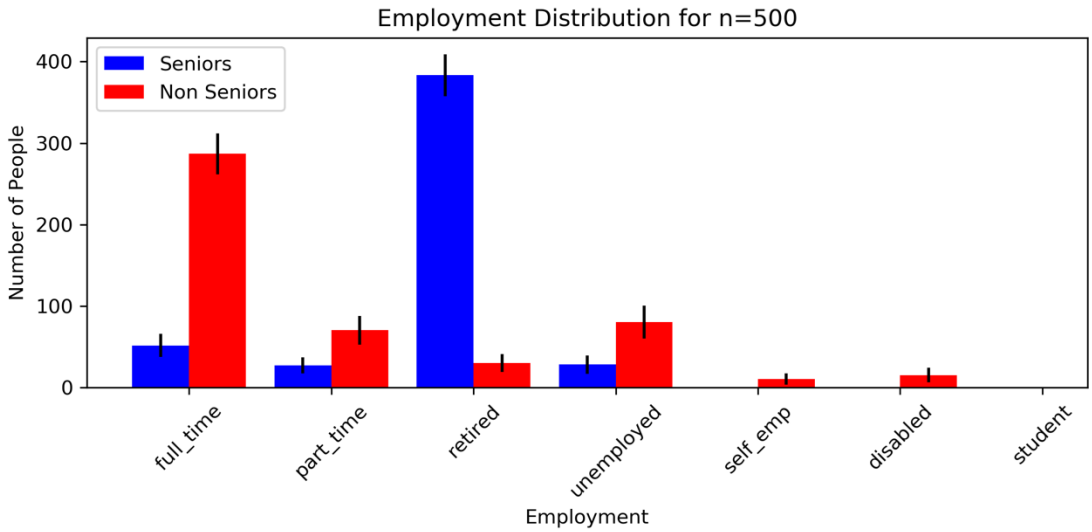
Lastly, seniors are slightly more prone than non-seniors to view the current high level of information availability as stressful rather than useful. This paints a promising picture overall for senior engagement with digital products such as online educational games.



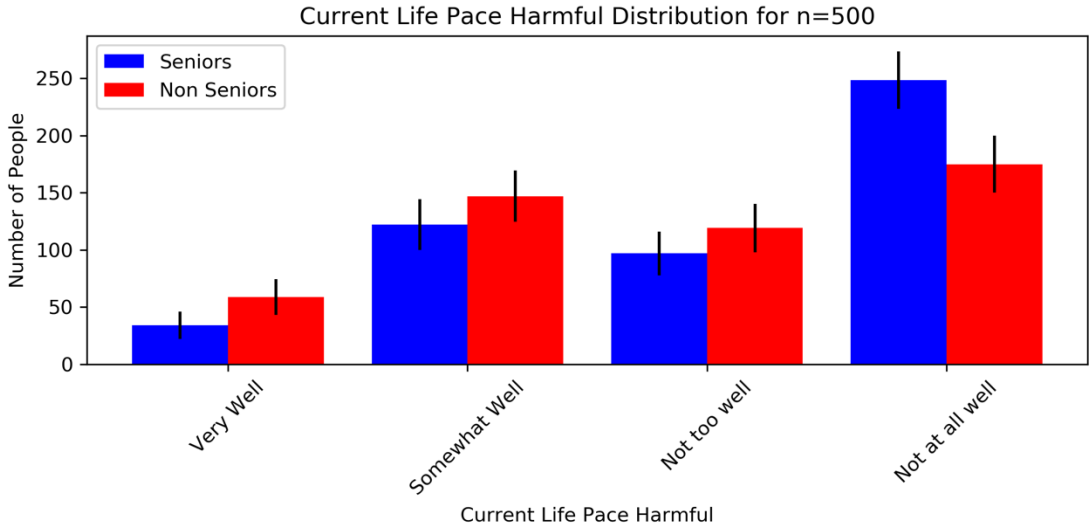
C) Time and Attention

We'll begin to make some comparisons between seniors and non-seniors by exploring the survey dataset.

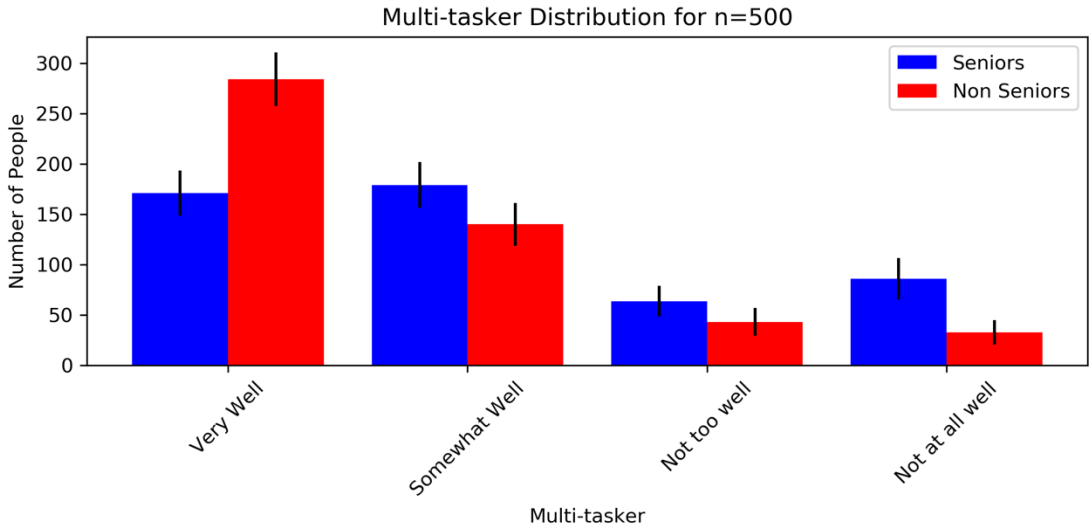
One of the best arguments for creating an educational product for seniors is the fact that they have free time because they are retired. Here's a quick sanity check:



But does retirement really mean that seniors have time and attention for a new activity in their life? Let's try to answer this indirectly. When asked whether they the idea that their "fast pace of life was harmful to their health", seniors were likely to disagree.

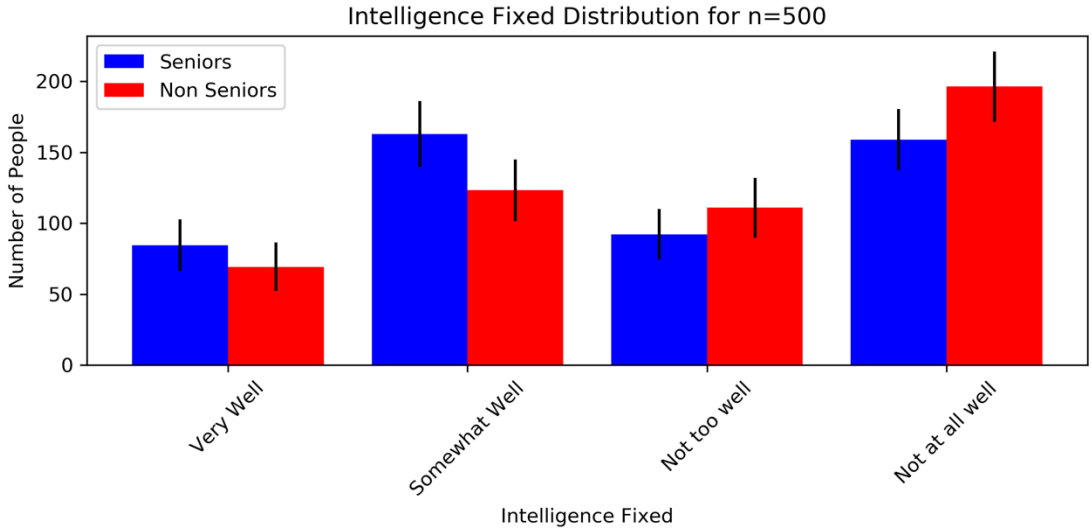


Seniors were also less likely to report that the idea that they were “usually doing two things at once” describes their views, which provides a second measure of “busy-ness”.

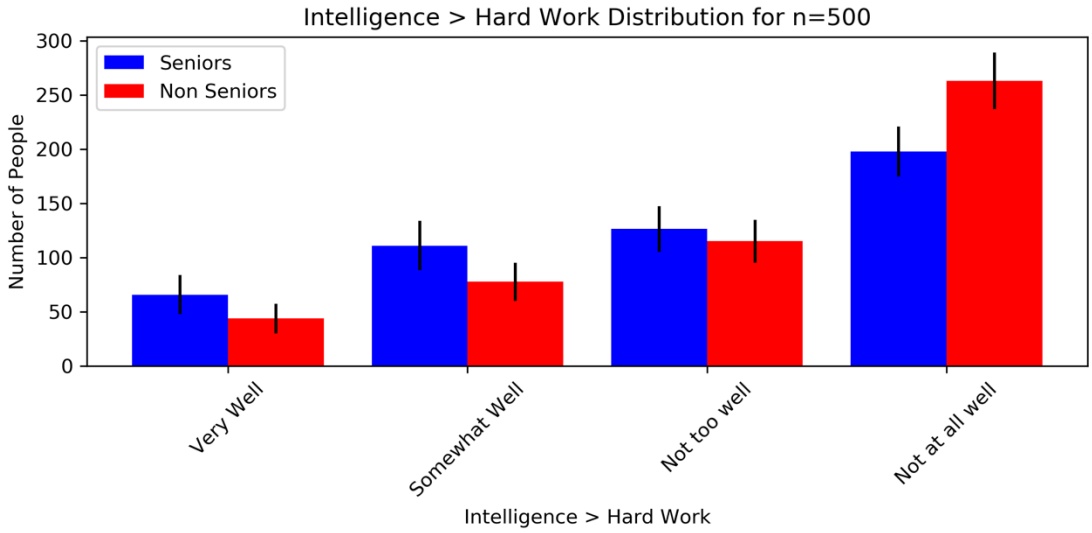


D) Growth Mindset

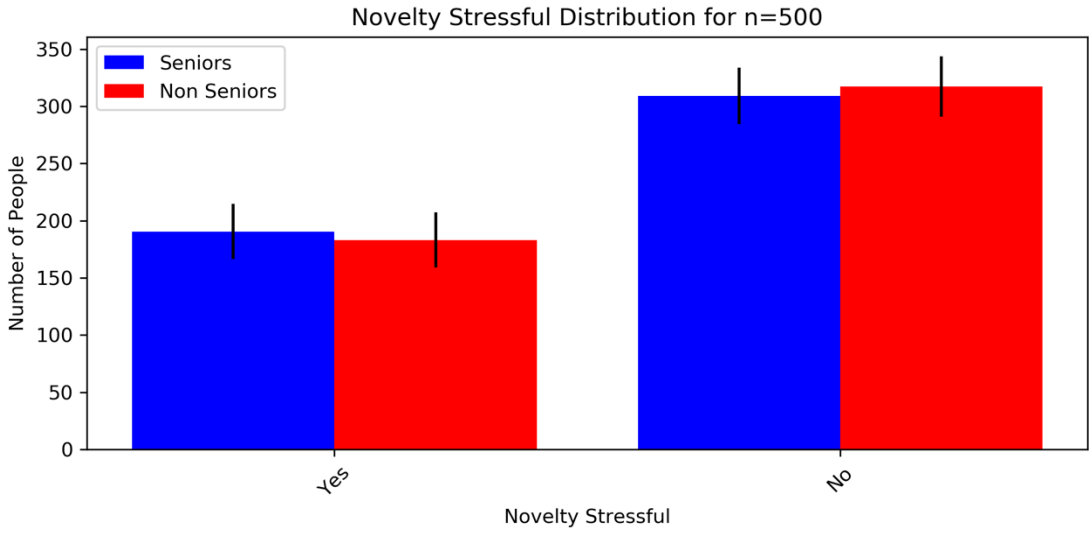
Overall, senior respondents demonstrated a slightly lower propensity towards growth mindset than non-seniors. For example, slightly more senior respondents reported that the statement “People can learn new things but cannot change their basic level of intelligence” described their views well. To some extent, we might expect this to be the case simply because growth mindset is a newer attitude towards learning whilst fixed talent/intelligence is an older one.



Senior respondents were slightly more likely to report that the statement “Truly smart people do not need to try hard” described their views well.



In terms of holistic openness to new experiences, seniors tended towards only slightly less preference for novelty. This bodes well for the game’s traction and provides a compelling counter-argument to those who might say seniors are globally unwilling to experiment. For example, senior respondents reported that trying new things was stressful only slightly more often than non-seniors.



For more detail on how graphical comparisons were made, see the “Exploratory Data Analysis for General Insight” section of the methodology notebook.