Data Analysis System Develop for physicians completing PANSS testing and the effect language has on the score

Kaisa Roggeveen, Scott Graham

Febuary 27th 2018

Summary

The intent of this report was to develop a data entry and analysis system for the PANSS instrument where results could be calculated in less than a half hour. The data entry system was designed to give each physician a unique ID, a language selection, a scale to input the severity of each symptom and save all entries into a file. The data analysis system was designed to retrieve the saved file and allow physicians to enter their unique ID. The data analysis system determined whether the physician received a passing score for each PANSS system and displayed the overall results for physicians receiving passing scores based on their language. From this analysis it was determined that language does influence the passing score of the physician.

Introduction

The Positive and Negative Syndrome Scale (PANSS) Instrument is a test used for accurately assessing the status of a patient's psychosis. The PANSS assessment involves physicians ranking a patient's psychological symptoms on a scale from 1 to 7, low to high respectively.

There are thirty different psychological symptoms that the physicians will rank. The first seven scaled symptoms are the positive symptoms, delusions, conceptual disorganization, hallucinatory behavior, excitement, grandiosity and auspiciousness/persecution and hostility. The next seven scaled symptoms are the negative symptoms, blunted affect, emotional withdrawal, poor rapport, passive/apathetic social withdrawal, difficulty in abstract thinking, lack of spontaneity and stereotyped thinking. The final symptoms scaled are the generic symptoms, somatic concern, anxiety, guilt feeling, tension, mannerisms and posturing, depression, motor retardation, cooperativeness, unusual thought content, disorientation, poor attention, lack of judgement and insight, disturbance of volition, poor impulse control, preoccupation and active social avoidance.

This report is studying the affects of the PANSS instrument being used in different languages and the affects that language has on a physician reeving a passing score. This report explains the development of the Apps for a PANSS testing workshop for the data entry and the data analysis.

The data entry app was designed to be used during the PANSS testing and the data analysis app was designed to be used after all physicians in the workshop had entered their data.

Data Collection Protocol

In order to collect the data, an application was developed using the Shiny Package in R. To use this application each physician must have access to a computer and the internet during the assessment.

The app is accessible through a webpage that is published by the Statistical Consulting Services from the Department of Mathematical Statistics at the University of Calgary. An example of this can be found at http://grahamst.at/shiny-server/PANSS/Input_App/.

This application gives each physician a unique Rater ID, and each physician would be able to select their choice of language to complete the ratings of the patients symptoms.

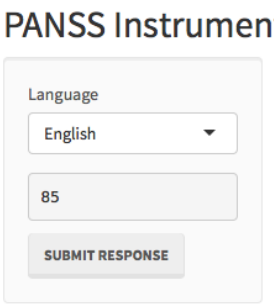


Figure 1: Image of the Language input

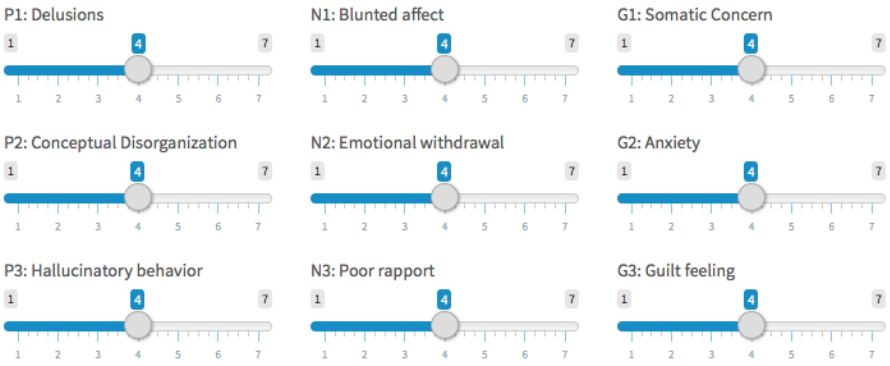


Figure 2: Image of the Slider-Scale

The application displays the three symptom sections, positive, negative and generic, in individual columns. The first column displays the seven positive symptoms, the second column displays the seven negative symptoms and the last column displays the 16 generic symptoms. Each physician can rate the symptom on a slider-scale between 1 (low) and 7 (high). Once the physician is satisfied with all of their responses, they can submit their responses. All of the submitted responses are collected in a comma-departed values (CSV) document or uploaded into a database table.

Data Analysis System

To create a quick and effective data analysis system an application was developed using the Shiny package in R. To view the results from the data analysis an individual physician would need access to a computer and the internet or the overall results could be shown to all physicians on an overhead screen.

This application is also accessible through a webpage that is published by the Statistical Consulting Services from the Department of Mathematical Statistics at the University of Calgary. An example of this can be found at http://grahamst.at/shiny-server/PANSS/Results_App/. Or seen in figures 3-5.

The application has a drop down menu where the physician can choose to display the results to a given question set, a check box set the results by language and the ability for physicians to enter their unique ID, and display their results.

The data analysis application gathers the data from the CSV file/database that was created in the input application. The data then runs through a data cleaner which checks that there are no responses outside of the scope of the questions.

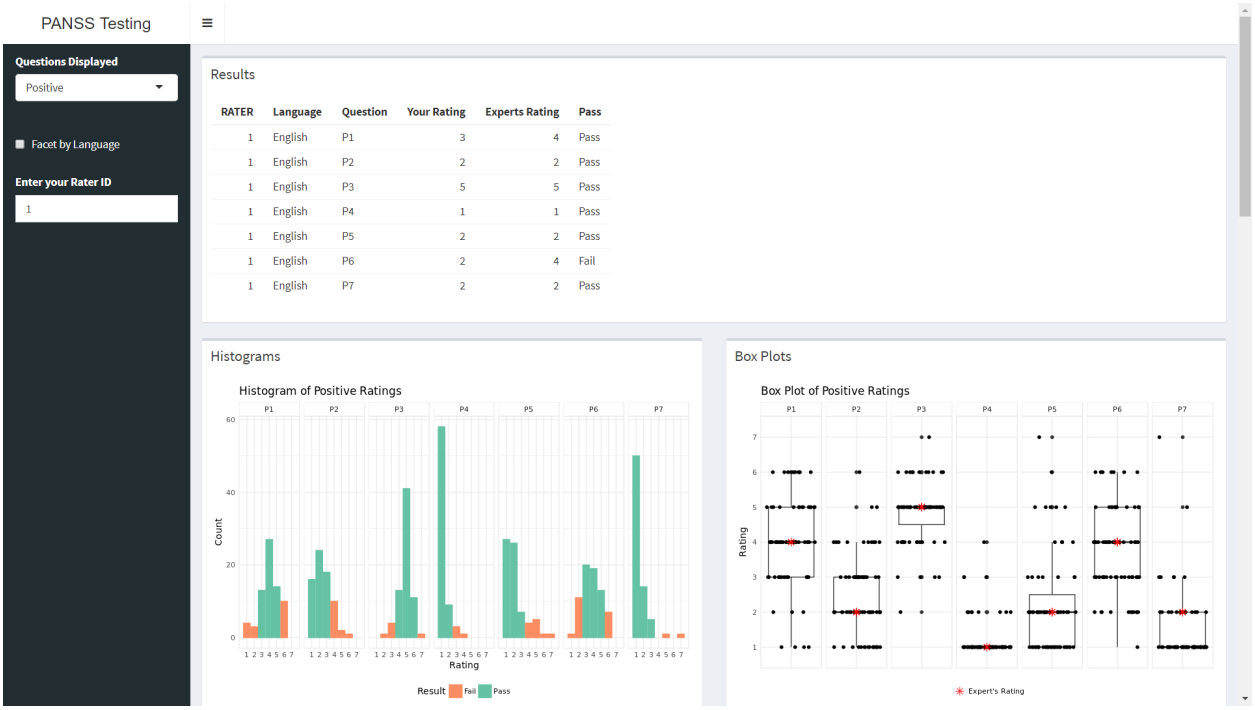


Figure 3: Results App Top

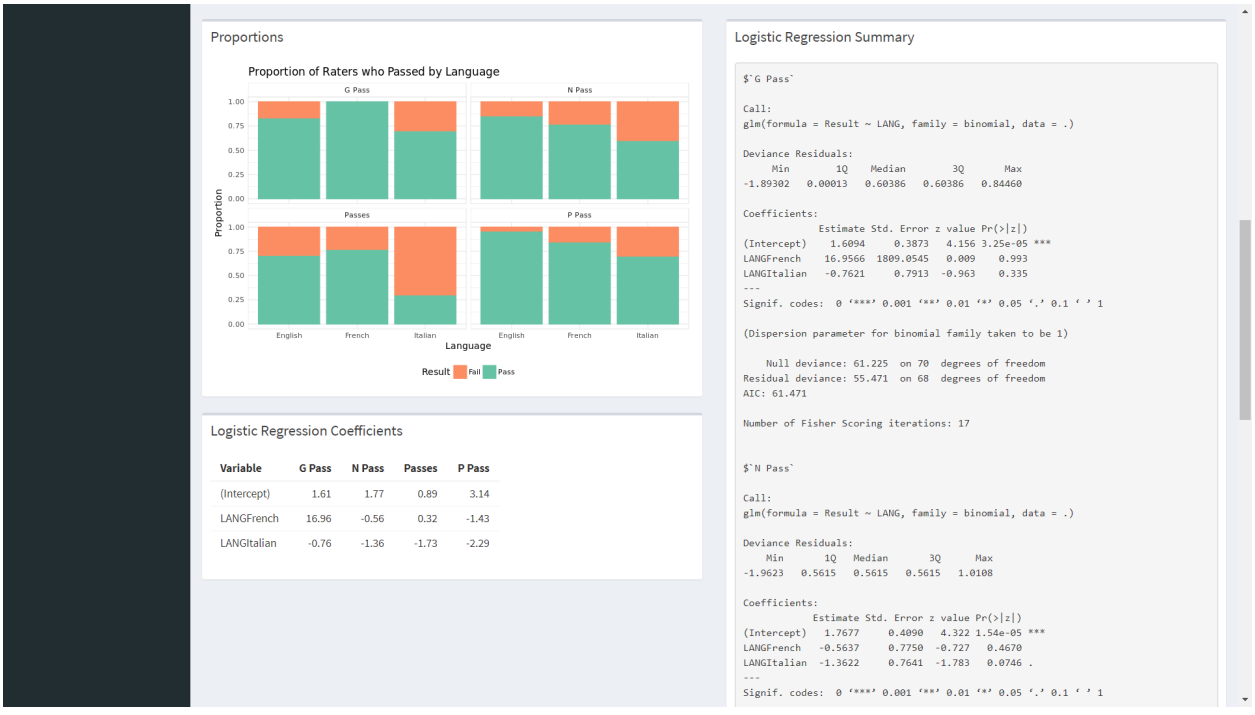


Figure 4: Results App Middle

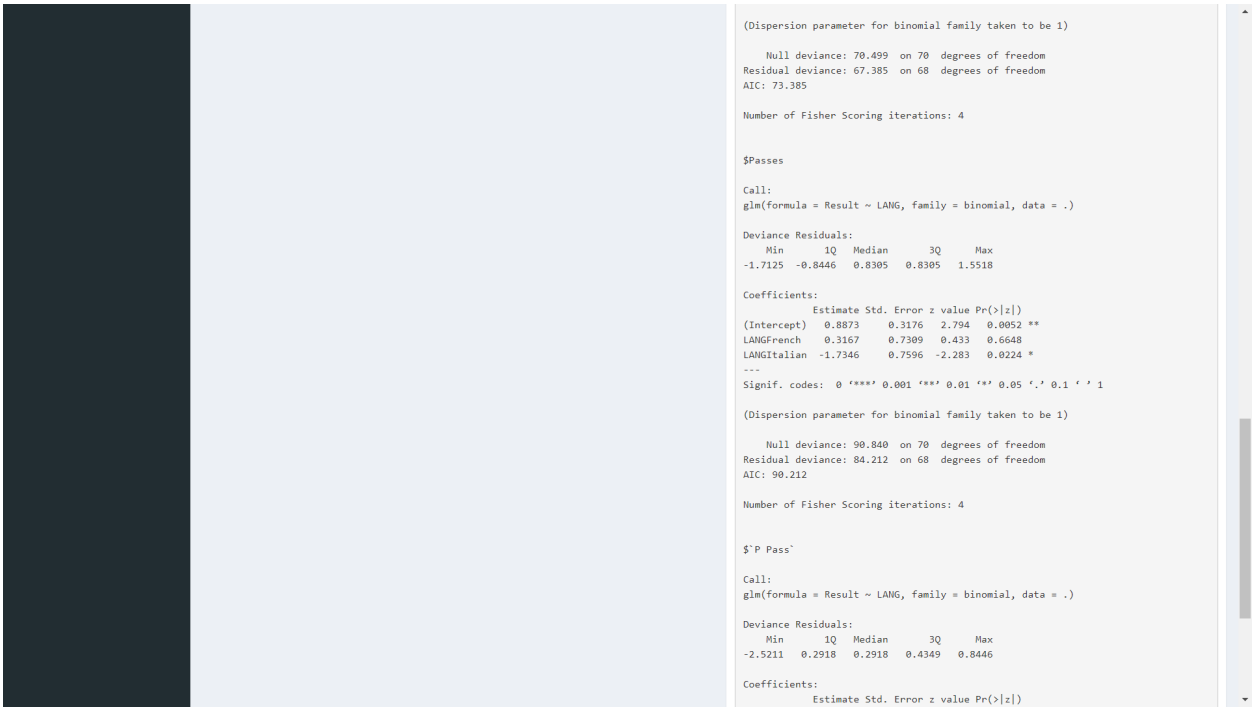


Figure 5: Results App Bottom

Data Analysis

Physician's Passing Score

To determine if a physician passed or failed the training they must meet certain criteria. The first criteria check was comparing each physician's symptom rating to an expert's symptom rating, if the physicians rating was within 1 of the expert's rating they received a pass on that symptom.

The second criteria check was the total number of "pass" questions in the positive, negative and generic symptoms sections. To receive a "pass" in the positive and negative symptoms sections at least 5 out of the 7 symptoms must be a "pass". To receive a "pass" in the generic symptoms section at least 10 out of the 16 must be a "pass".

The third criteria checked was if the physician "passed" all three sections. If the physician "passed" all three sections, then they received a "pass" for the PANSS training.

The following table shows the summary results of the physicians. The full results can bee seen in Appendix B.

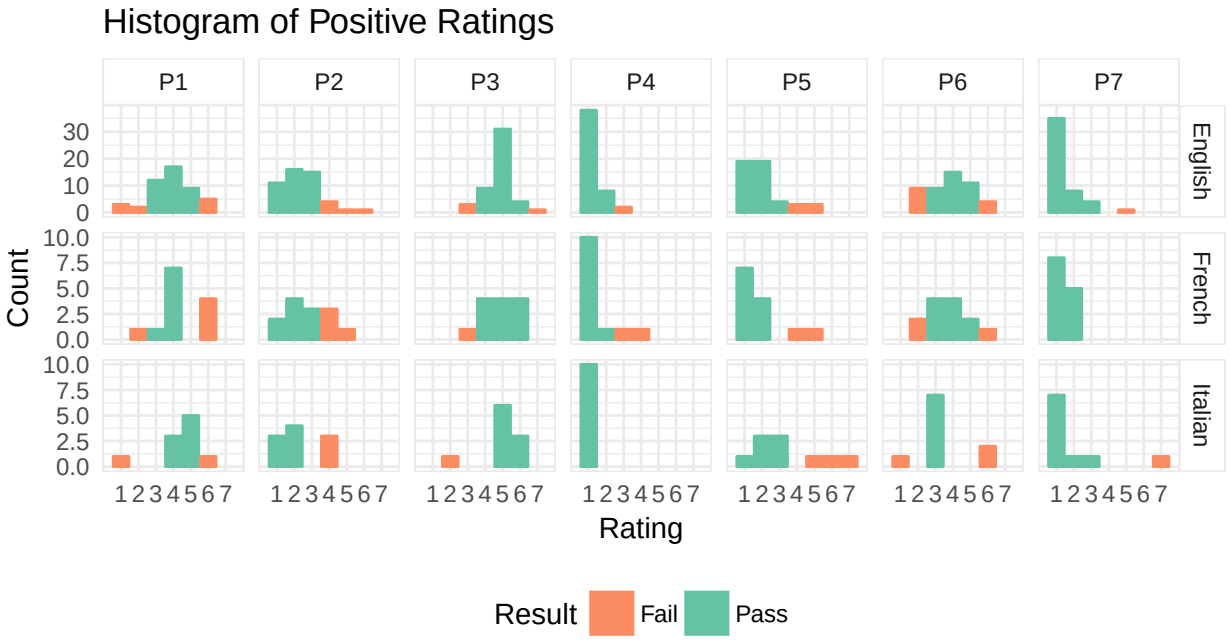
Table 1: Number of Passes

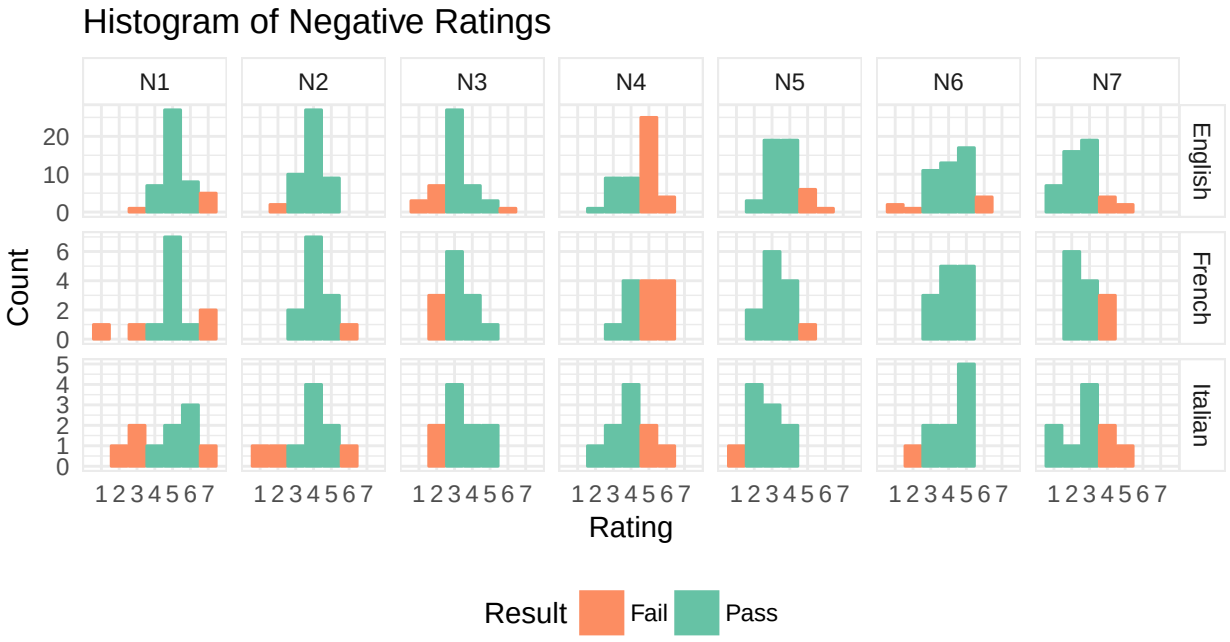
Language	Passed P	Passed N	Passed G	Passed	Total Physicians
English	46	41	40	34	48
French	11	10	13	10	13
Italian	7	6	7	3	10

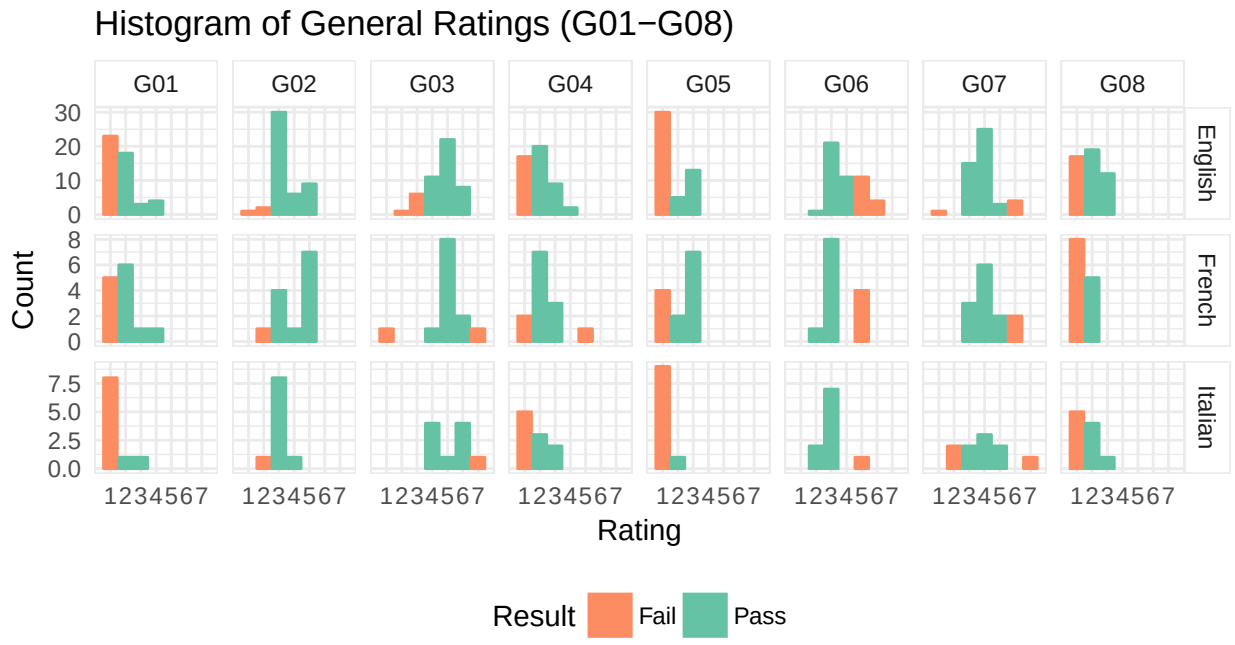
Table 2: Mean Number of Questions Passed

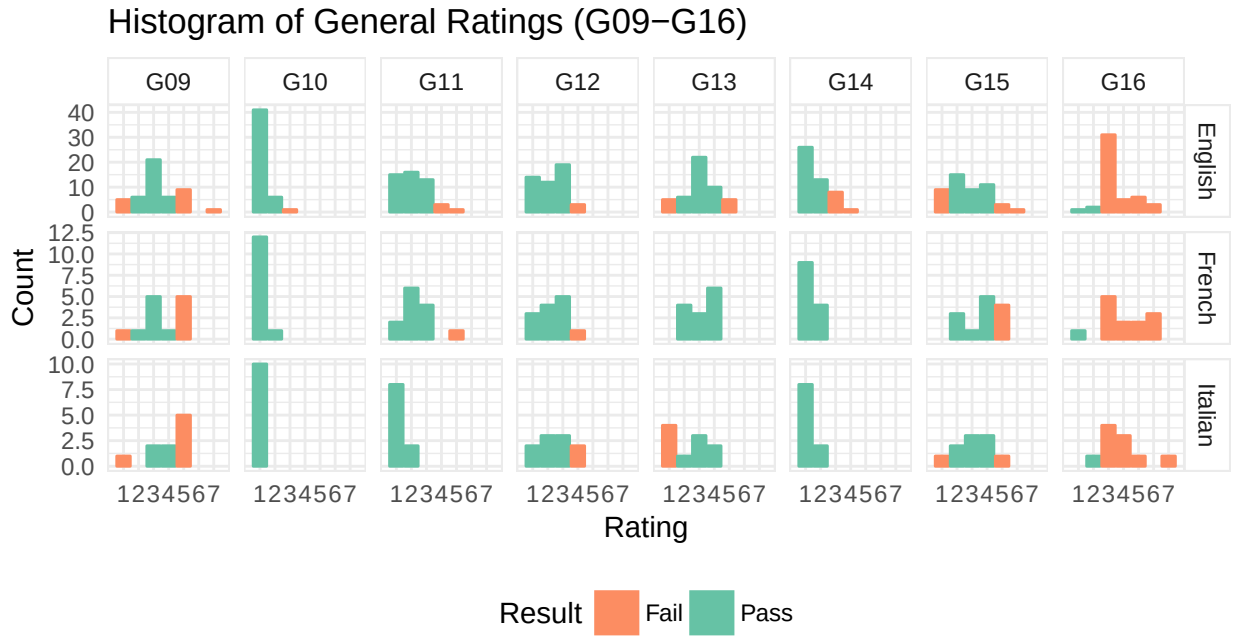
Language	P	N	G
English	6.1250	5.5833	11.4792
French	5.6923	5.4615	11.9231
Italian	5.7000	5.3000	10.4000

Table 3: SD Number of Questions Passed


Histogram of Positive Ratings
P1
P2
P3
P4
P5
P6
P7
English
French
Italian
Count
Rating
1234567
Result
Fail
Pass


Histogram of Negative Ratings
N1
N2
N3
N4
N5
N6
N7
English
French
Italian
Count
Rating
1234567
Result
Fail
Pass





Regression Analysis

We performed a logistic regression on the results from the 3 question sets, as well as the end result, using language as the predictor. For all 4 regressions, the English respondents are used as the base case for comparing against the French respondents and the Italian respondents.

	Estimate	Std. Error	z value	Pr(>|z|)

Table 4: Passes Question Set P by Language

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	3.135	0.7222	4.342	1.413e-05
LANGFrench	-1.431	1.055	-1.357	0.1749
LANGItalian	-2.288	0.9989	-2.291	0.02197

For the P questions, the odds in favor of passing of a French respondent compared to an English respondent increased on average by a multiplicative factor of 0.2391. For an Italian respondent compared to an English respondent, the odds in favor of passing increased on average by a multiplicative factor of 0.1014

Table 5: Passes Question Set N by Language

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.768	0.409	4.322	1.544e-05
LANGFrench	-0.5637	0.775	-0.7274	0.467
LANGItalian	-1.362	0.7641	-1.783	0.07464

For the N questions, the odds in favor of passing of a French respondent compared to an English respondent increased on average by a multiplicative factor of 0.5691. For an Italian respondent compared to an English respondent, the odds in favor of passing increased on average by a multiplicative factor of 0.2561

Table 6: Passes Question Set G by Language

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.609	0.3873	4.156	3.245e-05
LANGFrench	16.96	1809	0.009373	0.9925
LANGItalian	-0.7621	0.7913	-0.9631	0.3355

For the G questions, the odds in favor of passing of a French respondent compared to an English respondent can't be calculated as all of the French respondents passed the G questions, hence the nonsensical output in the regression. For an Italian respondent compared to an English respondent, the odds in favor of passing increased on average by a multiplicative factor of 0.4667

Table 7: Passes PANSS by Language

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	0.8873	0.3176	2.794	0.005203
LANGFrench	0.3167	0.7309	0.4333	0.6648
LANGItalian	-1.735	0.7596	-2.283	0.0224

For the overall test, the odds in favor of passing of a French respondent compared to an English respondent increased on average by a multiplicative factor of 1.3725. For an Italian respondent compared to an English respondent, the odds in favor of passing increased on average by a multiplicative factor of 0.1765

Conclusions

Recommendations

Appendix

Appendix A

```
print("foo")
```

Appendix B

Rater ID	Language	P Pass	N Pass	G Pass	Passes
1	English	TRUE	TRUE	TRUE	TRUE
2	French	TRUE	TRUE	TRUE	TRUE
3	English	TRUE	TRUE	FALSE	FALSE
4	English	TRUE	TRUE	FALSE	FALSE
5	English	TRUE	TRUE	TRUE	TRUE
6	English	TRUE	TRUE	TRUE	TRUE
7	English	TRUE	FALSE	TRUE	FALSE
8	English	TRUE	TRUE	TRUE	TRUE
9	English	TRUE	TRUE	TRUE	TRUE
10	English	FALSE	TRUE	TRUE	FALSE
11	English	TRUE	TRUE	TRUE	TRUE
12	Italian	TRUE	TRUE	TRUE	TRUE
13	English	TRUE	TRUE	TRUE	TRUE
14	English	FALSE	TRUE	FALSE	FALSE
15	English	TRUE	TRUE	TRUE	TRUE
16	French	TRUE	TRUE	TRUE	TRUE
17	French	TRUE	TRUE	TRUE	TRUE
18	English	TRUE	TRUE	FALSE	FALSE
19	English	TRUE	TRUE	TRUE	TRUE
20	English	TRUE	TRUE	TRUE	TRUE
21	English	TRUE	FALSE	FALSE	FALSE
22	English	TRUE	TRUE	TRUE	TRUE
23	English	TRUE	TRUE	TRUE	TRUE
24	English	TRUE	TRUE	TRUE	TRUE
25	English	TRUE	TRUE	TRUE	TRUE
26	English	TRUE	TRUE	TRUE	TRUE
27	Italian	FALSE	TRUE	TRUE	FALSE
28	English	TRUE	FALSE	TRUE	FALSE
29	English	TRUE	TRUE	TRUE	TRUE
30	French	TRUE	TRUE	TRUE	TRUE
31	Italian	TRUE	FALSE	FALSE	FALSE
33	English	TRUE	TRUE	TRUE	TRUE
34	French	TRUE	TRUE	TRUE	TRUE
35	French	TRUE	TRUE	TRUE	TRUE
36	English	TRUE	TRUE	TRUE	TRUE
37	English	TRUE	TRUE	TRUE	TRUE

Rater ID	Language	P Pass	N Pass	G Pass	Passes
38	Italian	TRUE	TRUE	TRUE	TRUE
39	English	TRUE	TRUE	TRUE	TRUE
40	Italian	FALSE	TRUE	TRUE	FALSE
41	English	TRUE	TRUE	TRUE	TRUE
42	French	TRUE	FALSE	TRUE	FALSE
43	Italian	TRUE	TRUE	TRUE	TRUE
44	French	FALSE	FALSE	TRUE	FALSE
45	French	TRUE	TRUE	TRUE	TRUE
46	French	TRUE	TRUE	TRUE	TRUE
47	English	TRUE	TRUE	TRUE	TRUE
48	English	TRUE	FALSE	TRUE	FALSE
49	English	TRUE	TRUE	TRUE	TRUE
50	English	TRUE	TRUE	TRUE	TRUE
51	French	TRUE	TRUE	TRUE	TRUE
52	English	TRUE	TRUE	TRUE	TRUE
53	English	TRUE	TRUE	TRUE	TRUE
54	English	TRUE	TRUE	TRUE	TRUE
55	Italian	FALSE	TRUE	FALSE	FALSE
56	English	TRUE	TRUE	TRUE	TRUE
57	Italian	TRUE	FALSE	TRUE	FALSE
58	English	TRUE	FALSE	FALSE	FALSE
59	English	TRUE	TRUE	TRUE	TRUE
60	French	FALSE	FALSE	TRUE	FALSE
61	Italian	TRUE	FALSE	FALSE	FALSE
62	Italian	TRUE	FALSE	TRUE	FALSE
63	English	TRUE	FALSE	TRUE	FALSE
64	English	TRUE	TRUE	TRUE	TRUE
65	English	TRUE	TRUE	TRUE	TRUE
66	English	TRUE	TRUE	TRUE	TRUE
67	English	TRUE	FALSE	TRUE	FALSE
68	French	TRUE	TRUE	TRUE	TRUE
69	English	TRUE	TRUE	FALSE	FALSE
70	English	TRUE	TRUE	FALSE	FALSE
71	English	TRUE	TRUE	TRUE	TRUE
72	English	TRUE	TRUE	TRUE	TRUE