Data analysis and statistics course

End of semester assignment

Name:

Student number

Date submitted

Table of Contents

[1 Question 1 - 35 marks 1](#_Toc104462940)

[1.1 Explore the data descriptively, creating appropriate tables and figures, where needed. 1](#_Toc104462941)

[1.2 Mixed Linear Model analyses 1](#_Toc104462942)

[2 Question 2 - 20 marks 2](#_Toc104462943)

[2.1 Descriptive results 2](#_Toc104462944)

[2.2 Compute measures of internal consistency for each of the scales. 4](#_Toc104462945)

[2.2.1 Reliability analysis 4](#_Toc104462946)

[2.3 Reliability analysis: Failure to fail 5](#_Toc104462947)

[3 Question 3 - 45 marks 7](#_Toc104462948)

[3.1 Model 1: Resilience and Fear of Failure effects on outcomes (reading fluency) 8](#_Toc104462949)

[4 Model 2: Resilience and Failure effects on outcomes (reading fluency) 10](#_Toc104462950)

# Question 1 - 35 marks

**Introduction**

This section provides results on analysis of interventions that try to promote growth mindsets in children learning to read. There are two groups in the analysis, namely a control group (control), and an intervention group (growth, or growth\_mindset). The control group receives the usual classroom activities, whereas the growth mindset group spends an hour each week of the year doing activities aimed at promoting a growth mindset. Each child is tested at the beginning of the program (January), and then halfway through the program (June), and again at the end of the program (December). The analysis will provide descriptive statistics of the results between the 2 groups.

## Explore the data descriptively, creating appropriate tables and figures, where needed.

The following table provides descriptive statistics of the two groups regards reading scores, The mean is the arithmetic average while the standard deviation measures how far the results deviate from the mean. Notably, the standard deviation is larger than 2, showing that the score range was large: minimum was 7 and maximum 32 for the entire group.

**Table 1:** Descriptive statistics between groups

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Obs | Mean | Std.dev | Min | Max |
| All | 29 | 16.55172 | 6.982561 | 7 | 32 |
| Control | 15 | 14.33333 | 5.037951 | 7 | 22 |
| Growth | 14 | 18.92857 | 8.118964 | 9 | 32 |

The mean reading score fora control group (control) (M=14.3, SD=5) was lower compared to the intervention group (growth) (M=18, SD=8). The maximum reading score for the control group (22) was much lower than that for the growth group (M=32). The performance on a standardized reading test, maximum score of 35.

## Mixed Linear Model analyses

**Hypothesis:** The growth group will improve more than the control group does over time; the researcher thinks that both groups will improve over time.

**Table 2:** Mixed Linear regression model to show the differences in scores between control and intervention group

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | SS | df | MS | Number of obs | = | 29 |
|  |  |  |  | F(2, 26) | = | 4.25 |
| Model | 336.367382 | 2 | 168.183691 | Prob > F | = | 0.0253 |
| Residual | 1028.80503 | 26 | 39.5694243 | R-squared | = | 0.2464 |
|  |  |  |  | Adj R-squared | = | 0.1884 |
| Total | 1365.17241 | 28 | 48.7561576 | Root MSE | = | 6.2904 |

The table above shows pre-model statistics. Firstly, the model is fit, as the F-statistics is significant at 5% level. The R-squared, which is the explanatory power of the model is 0.24, meaning that the model explains the variance between the intervention and control group 24 of the times. This is low explanatory power, caused by few variables in the model. To enhance the model, there is need for other variables such as age, gender, and grade of the students.

**Table 3:** The model estimates

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| reading\_sc~e | Coefficient | Std. err. | t | P>t | [95% conf. | interval] |
| Month (Dec) | -3.113208 | 1.445842 | -2.15 | 0.041 | -6.085179 | -.141236 |
| Treat |  |  |  |  |  |  |
| Growth | 4.81761 | 2.339875 | 2.06 | 0.050 | .0079283 | 9.627292 |
| \_cons | 20.55975 | 3.316595 | 6.20 | 0.000 | 13.74239 | 27.37711 |

The table above shows that the treatment group (intervention) was 4.8 times likely to improve more than the control group does over time (β=4.8, p<0.1). Though weak evidence, the results are significant at 10%, hence the null hypothesis is rejected, and a conclusion is made that both groups will improve over time.

# Question 2 - 20 marks

Two scales that were measured in the PISA 2018 data were those of *resilience*, and *fear of failure*. These are referenced in the codebook referred to above.

Combine these items in a factor analysis and assess whether there is good evidence for a two-factor structure, each factor corresponding to one of the original scales. Compute measures of internal consistency for each of the scales.

## Descriptive results

Before running a regression analysis, it is important to explore the distribution of the data, through mean scores, and how far the average score deviates from the norm (standard deviation). In this section we run descriptive summaries of the Likert scale responses.

**Table 2.1**: Scale measuring resilience

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Obs | Mean | Std. dev. | Min | Max |
| ST188Q01HA | 480 | 3.10625 | .4996476 | 1 | 4 |
| ST188Q02HA | 480 | 3.23125 | .6118715 | 1 | 4 |
| ST188Q03HA | 479 | 2.778706 | .7550982 | 1 | 4 |
| ST188Q06HA | 479 | 2.759916 | .7966856 | 1 | 4 |
| ST188Q07HA | 478 | 3.031381 | .6117804 | 1 | 4 |

Table 2.1 indicates that the 5 items that the five items in the Resilience construct had mean scores of around 3 (agreement), overall, they felt they had resilience. They agreed that they usually manage one way or another (M=3.1; SD=0.49), that they feel proud that they have accomplished things (M=3.2; SD=0.61), and that they can usually find their way out of it (M=3.0; SD=0.611). They weakly agreed that they feel that they can handle many things at a time (M=2.8; SD=0.76) and that their belief in themselves gets them through hard times (M=2.7; SD=0.80). These statements were combined to create the index of self-efficacy (RESILIENCE), depicted the chart that follows.

Chart, histogram

Description automatically generated

**Figure 1:** Histogram showing the Likert scale responses-Resilience constructs

The scores computed from all questions relating to the children’s resilience were negatively skewed also, with a mean of around 2.9, indicating they generally agreed that they were resilient. These positive values in this index mean that the student reported higher self-efficacy than did the average student across OECD countries.

**Table 2.2**: Scale measuring fear of failure

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Obs | Mean | Std. dev. | Min | Max |
| ST183Q01HA | 484 | 2.766529 | .9265106 | 1 | 4 |
| ST183Q02HA | 485 | 2.723711 | .9033653 | 1 | 4 |
| ST183Q03HA | 484 | 2.828512 | .9421865 | 1 | 4 |

Table xx indicates that the respondents weakly agreed with all aspects of this scale, compared to resilience. They weakly agreed that when failing, they worry about what others think of them (M=2.7; SD=0.9); they were afraid that they might not have enough talent (M=2.7; SD=0.92); and that sometimes they doubted their plans (M=2.8; SD=0.94). These statements were combined to create the index of fear of failure (GFOFAIL), depicted in the following graph.

Chart, histogram

Description automatically generated

**Figure 2:** Histogram showing the Likert scale responses-FEAR constructs

Figure 2 (GFOFAIL) indicates that the scores for this construct were not normally distributed, more values lie to the left, hence negatively skewed. The mean lies around 2.7, indicating they generally agreed that they had fear of failure. These positive values in this index mean that the student expressed less fear of failure than did the average student across OECD countries.

## Compute measures of internal consistency for each of the scales.

Cronbach’s alpha reliability coefficient normally ranges between 0 and 1. The closer Cronbach’s alpha coefficient is to 1.0 the greater the internal consistency of the items in the scale. George and Mallery (2003) provide the following rules of thumb: *“\_ > .9 – Excellent, \_ > .8 – Good, \_ > .7 – Acceptable, \_ > .6 – Questionable, \_ > .5 – Poor, and\_ < .5 – Unacceptable*” (George and Mallery, 2018. pp249-260). All items should correlate the total score, so we look for items that don’t correlate with the overall score from the scale.

### Reliability analysis

Call: alpha(x = computeResilience)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | raw\_alpha | std.alpha | G6(smc) | average\_r | S/N | ase | mean | sd | median\_r |
|  | 0.42 | 0.3 | 0.46 | 0.12 | 0.43 | 0.038 | 2.8 | 0.57 | -0.12 |

The raw alpha shows Cronbach’s α for the overall computed scale (Resilience) of 0.42, which is poor. This could be due to the fact that there were few items (3 questions) in the scale.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Reliability if an item is dropped: | | |  |  |  |  |  |  |
|  | raw\_alpha | std.alpha | G6(smc) | average\_r | S/N | alpha se | var.r | med.r |
| ST188Q07HA | 0.81 | 0.81 | 0.68 | 0.68 | 4.26 | 0.017 | NA | 0.68 |
| ST183Q01HA | -0.43 | -0.47 | -0.19 | -0.19 | -0.32 | 0.118 | NA | -0.19 |
| ST183Q02HA | -0.24 | -0.27 | -0.12 | -0.12 | -0.21 | 0.102 | NA | -0.12 |

The individual Cronbach’s alpha values for individual items show two items whose reliability was good (ST188Q01HA) and (ST188Q02HA), while the rest had low reliability, but not bad. This means that the internal consistency of questions was good for the two questions, they were understood by the respondents more than other questions.

## Reliability analysis: Failure to fail

Call: alpha(x = computeGfofail)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| raw\_alpha | std.alpha | G6(smc) | average\_r S/N ase | mean sd | median\_r |
| 0.00022 | 0.61 | 0.6 | 0.17 1.5 0.00078 | 170 123 | 0.14 |

The raw alpha shows a Cronbach’s of α for the overall computed scale (GFOFAIL) of 0.61, which is questionable. This could be due to the that most respondents were biased toward agreement.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Reliability if an item is dropped | | | | | | | | |
|  | raw\_alpha | **std.alpha** | G6(smc) | average\_r | S/N | alpha | se | var.r |
| ST188Q01HA | 2.8e-04 | **0.6651** | 0.53 | 0.16 | 1.15 | 0.00073 | 0.026 | 0.11 |
| ST188Q02HA | 1.4e-04 | **0.6645** | 0.54 | 0.16 | 1.12 | 0.00070 | 0.029 | 0.11 |
| ST188Q03HA | 2.5e-05 | **0.6291** | 0.53 | 0.15 | 1.09 | 0.00068 | 0.023 | 0.14 |
| ST188Q06HA | 4.9e-05 | **0.6272** | 0.52 | 0.15 | 1.08 | 0.00066 | 0.023 | 0.14 |
| ST188Q07HA | 3.8e-04 | **0.5869** | 0.48 | 0.14 | 0.97 | 0.00068 | 0.017 | 0.14 |

The individual Cronbach’s alpha values for individual items show two items whose reliability was good (ST188Q01HA) and (ST188Q02HA), while the rest had low reliability, but not bad. This means that the internal consistency of questions was good for the two questions, they were understood by the respondents more than other questions. The raw alpha values for the three questions are lower than the overall scale reliability, in such situations, it is advisable to drop these particular items so as to increase the overall α of the scale.

**Factor analysis**

It should also be noted that while a high value for Cronbach’s alpha indicates good internal consistency of the items in the scale, it does not mean that the scale is unidimensional. Factor analysis is a method to determine the dimensionality of a scale, in this case a Principal Components Analysis (PCA) will be applied. It uses algorithms to "reduce" data into correlated "factors" that provide a conceptual and mathematical understanding of the construct of interest[[1]](#footnote-1). Under the assumption that researchers are measuring for one construct, the individual items should correlate in some form or fashion (Jones et al., 2018). Principal Components Analysis (PCA) reduces survey data down into content areas that account for the most variance (Muka et al., 2013).

**Factor analysis: Fear of failure**

Table xx: Principal components/correlation

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Component | Eigenvalue | Difference | Proportion | Cumulative | |
| Comp1 | 2.25945 | 1.38006 | 0.4519 | 0.4519 |  |
| Comp2 | .879388 | .109579 | 0.1759 | 0.6278 |  |
| Comp3 | .769809 | .132936 | 0.1540 | 0.7817 |  |
| Comp4 | .636872 | .18239 | 0.1274 | 0.9091 |  |
| Comp5 | .454482 | . | 0.0909 | 1.0000 |  |

The table above are the loadings showing the proportion of variance explained by each factor. The table shows that there is one factor that is extracted in this model, it has eigenvalue of 1 and contributes to 45.2% of variance in the fear to fail scale. The following table provides which actual aspect has much influence in this scale.

**Table xx:** Principal components (eigenvectors)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Comp1 | Comp2 | Comp3 | Comp4 | Comp5 |
| ST188Q01HA | 0.4066 | 0.6621 | -0.4369 | -0.0597 | 0.4493 |
| ST188Q02HA | 0.3907 | 0.4355 | **0.7747** | 0.0346 | -0.2374 |
| ST188Q03HA | 0.4585 | -0.2975 | -0.0857 | 0.8315 | 0.0506 |
| ST188Q06HA | 0.4578 | -0.5163 | 0.2384 | -0.4442 | 0.5192 |
| ST188Q07HA | 0.5120 | -0.1300 | -0.3806 | -0.3264 | -0.6852 |

The factor in the FEAR scale that contributes to twice as much variance in the scale is ST188Q02HA (I feel proud that I have accomplished things). It has the highest principal component loadings in this scale. The implication of the results is that the children felt very proud when they managed to read fluently-this is what they valued most in this scale.

**Factor analysis: Resilience**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Component | Eigenvalue | Difference | Proportion | Cumulative |
| Comp1 | 2.25047 | 1.7785 | 0.7502 | 0.7502 |
| Comp2 | .471971 | .194414 | 0.1573 | **0.9075** |
| Comp3 | .277557 | . | 0.0925 | 1.0000 |

The table above are the loadings showing the proportion of variance explained by each factor. The table shows that there is one factor that is extracted in this model, it has eigenvalue of 1 and contributes to 75.2% of variance in the fear to fail scale. The following table provides which actual aspect has much influence in this scale

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Comp1 | Comp2 | Comp3 |
| ST183Q01HA | 0.5667 | -0.6759 | 0.4711 |
| ST183Q02HA | 0.6049 | -0.0468 | -**0.7949** |
| ST183Q03HA | 0.5594 | 0.7355 | 0.3824 |

The factor in the FEAR scale that contributes to twice as much variance in the scale is ST138Q02HA (When I am failing, I am afraid that I might not have enough talent). It has the highest principal component loadings in this scale. The implication of the results is that the children felt their failure was due to their lack of reading talent.

# Question 3 - 45 marks

A key outcome measure in the PISA 2018 dataset, for our purposes, is reading fluency. Build, test, and report a regression model that takes reading fluency as the outcome variable, and models it with predictors available to you in the dataset. Select variables for inclusion in the model on the basis of whether they seem likely to you to be useful predictors. Remember to think about possible control variables (and how you would treat them as controls). **Provide useful tables and figures summarizing your analysis, and findings.**

**Variables in the equation**

1. Reading fluency: A total score created expediently for this question paper (but not strictly correctly) from the original data, reflecting each student's level of reading fluency
2. GFORFAIL: Fear of failure
3. Resilience:

## Model 1: Resilience and Fear of Failure effects on outcomes (reading fluency)

In linear regression, the variable to be predicted is called the *dependent variable* (or sometimes, the outcome variable). The variable used to predict the other variable's value is called the *independent variable* (or sometimes, the predictor variable).

1. **Dependent variable**
   1. **Reading fluency:** A total score created expediently for this question paper (but not strictly correctly) from the original data, reflecting each student's level of reading fluency.
2. **Independent variables:** 
   1. **Measure of resilience:** PISA 2018 asked (ST188) students to report the extent to which they agreed with the following statements about themselves: *“I usually manage one way or another”; “I feel proud that I have accomplished things”; “I feel that I can handle many things at a time”; “My belief in myself gets me through hard times”; and “When I’m in a difficult situation, I can usually find my way out of it”.* These statements were combined to create the index of self-efficacy (RESILIENCE). Positive values in this index mean that the student reported higher self-efficacy than did the average student across OECD countries.
   2. **Fear of failure:** Students in PISA 2018 were asked to report the extent to which they agree ( “strongly disagree”, “disagree”, “agree”, “strongly agree”) with the following statements (ST183): “When I am failing, I worry about what others think of me”; “When I am failing, I am afraid that I might not have enough talent”; and “When I am failing, this makes me doubt my plans for the future”. These statements were combined to create the index of fear of failure (GFOFAIL). Positive values in this index mean that the student expressed greater fear of failure than did the average student across OECD countries.

**Research question: Do fear of results and resilience predict reading fluency in this study?**

**Call:**

*lm(formula = reading\_fluency ~ GFOFAIL + RESILIENCE\_1, data = PISA\_2018)*

The call section shows the formula that R used to fit the regression model. Reading fluency is dependent variable and Fear of failure (GFOFAIL) and Resilience are the predictors.

*Residuals:*

Min 1Q Median 3Q Max

-16.0566 -1.0683 0.7867 1.8457 3.2972

The residuals are the difference between the actual values and the predicted values. The distribution of the data parts much from normality (is not symmetrical), as the median is near 1, implying that the model does not predict the outcomes so well. Hence, the results should be interpreted with caution.

*Coefficients:*

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

(Intercept) 13.7073 2.5602 5.354 1.34e-07 \*\*\*

GFOFAIL -0.5760 0.7776 -0.741 0.459

RESILIENCE 0.9047 0.9853 0.918 0.359

Signif. codes: ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.727 on 480 degrees of freedom

(17 observations deleted due to missingness)

Multiple R-squared: 0.004043, Adjusted R-squared: -0.0001069

F-statistic: 0.9742 on 2 and 480 DF, p-value: 0.3782

Estimates:

1. *Fear of failure:* The results indicate that there was no statistically significant association between fear of failure and reading fluency (B=-0.576, p>0.05). However, the results provide an insight into that there is a potential negative relationship between fear of failure and reading fluency, implying that as fear increases reading fluency drops. These results are supported by some literature that states that [add context from a psychological lens]
2. *Resilience:* The results indicate that there was no statistically significant association between resilience score and reading fluency (B=-0.905, p>0.05). However, the results provide an insight into that there is a potential positive relationship between resilience and reading fluency, implying that as resilience increased the reading fluency improved. In other words, the student who reported higher self-efficacy were likely to report high reading fluency in this study. Descriptive results showed the students expressed lesser fear of failure than did the average student across OECD countries. These results are supported by some literature that states that [add context from a psychological lens]

*Standard error:* The standard error of the coefficient is an estimate of the standard deviation of the coefficient. In effect, it is telling how much uncertainty there is with the coefficient.

*The Multiple R-squared*: It is the percentage of the variation within the dependent variable that the independent variable is explaining, in other words it is the explanatory power of the model, in this case it is too low, hence the model does not perfectly predict the outcomes (reading fluency). There is need to add more variables in the equation to increase the model explanatory power.

# Model 2: Resilience and Failure effects on outcomes (reading fluency)

HEDRES + PARED + SCREADCOMP + SCREADDIFF + JOYREAD+ STIMREAD + TMINS

(optional)

REFERENCES

George, D. and Mallery, P., 2018. Reliability analysis. In IBM SPSS Statistics 25 Step by Step (pp. 249-260). Routledge.

Muca, M., Puka, L., Bani, K. and Shahu, E., 2013. Principal components and the maximum likelihood methods as tools to analyze large data with a psychological testing example. European Scientific Journal, 9(20).

Jones, P.J., Mair, P. and McNally, R.J., 2018. Visualizing psychological networks: A tutorial in R. Frontiers in Psychology, p.1742.

1. <https://www.scalestatistics.com/principal-components-analysis.html> [↑](#footnote-ref-1)