**Contents**

**Research Questions & Hypotheses**

The purpose of the research study is to identify determinants of Ph.D. completion among candidates including GPA and other selector metrics. To fulfill the purpose of the study, the following research questions have been identified,

1. What are the determinants of Ph.D. completion among successful candidates?

2. Is GPA a proper factor to determine Ph.D. completion among candidates?

Based on the above research questions, the following Hypotheses have been devised,

H1 (Hypothesis 1): GPA, GRE determines Ph.D. completion among candidates.

H2 (Hypothesis 2): Letter of Recommendations, Student Motivation, determines Ph.D. completion among candidates.

H3 (Hypothesis 3): Age, Gender determines Ph.D. completion among students.

H4 (Hypothesis 4): Emotional Stability, Financial Resources, Hostility and Social Abilities Mean Rating of Selectors Impression of Applicant determines Ph.D. completion among candidates.

**Introduction & Methodology**

**Research Design**

The methodology of this research study is based upon a quantitative research design based on a critical realist approach. Since the research study considers a sample in real-time, therefore, a realist approach has been opted for, explaining the determinants of Ph.D. completion among candidates. Moreover, based on the approach, the study can be called deductive since it deduces generalizations based on results obtained through a proper detailed analysis provided in the

report ahead. On the other hand, as mentioned before the research strategy that has been opted within this study is quantitative, using a more descriptive approach so that the determinants or predictors of Ph.D. completion among candidates can be discussed in detail. Instead of a longitudinal research study, a cross-sectional study has been considered to avoid any discrepancies such as differences in grading plans, etc.

For the data collection process, among the sample chosen, a survey was distributed which was filled by the students gradually. Therefore, a structured questionnaire was used for the survey. The survey was based on 18 variables including gender, age, GPA, GRE scores, letter of recommendations, motivation, stability, financial funding, marital status, age, social skills, hostility, and impression. Among all these variables Ph.D. completion is the dependent variable whereas GPA, GRE Scores, Motivation, Stability, Financial Resources, Hostility, Impression, Letter of Recommendation, and Social Abilities have been identified to be independent variables. All these variables have been measured on a 9-point Hedonic scale ranging from extremely low to extremely high.

**Sample**

Since the study is regarding Ph.D. completion, 100 Ph.D. candidates from higher degree institutes have been selected for this study's sample. Each candidate was asked regarding the status of completion of their degrees whereas the entire sample was chosen based on convenient sampling. Convenient sampling helped in supporting the researcher since it provided direct ease in selecting candidates who were easily in reach and available, reducing time and cost.

**Buisness Understanding**

For Data Analysis, the multiple linear regression model has been utilized using the SPSS software to avoid any kinds of human error within the analysis of the results. As mentioned above, Ph.D. completion has been considered as the dependent variable GPA, GRE Scores, Gender, Age at entry ,Motivation, Stability, Financial Resources, Hostility, Impression, Letter of Recommendation, and Social Abilities have been identified to be independent variables.

<div align="center">

**Results**

</div>

**Descriptive Statistics**

Around 100 Ph.D. candidates were engaged within this research study to provide an effective representation of the small sample considered. Among the entire sample, there were around 64% Females i.e., 64 and 36% Males i.e., 36 out of a sample of 100 as displayed in Table 1.

Table 1. Gender

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| | | | | | |
| | Female | 64 | 64.0 | 64.0 | 64.0 |
| Valid | Male | 36 | 36.0 | 36.0 | 100.0 |
| | Total | 100 | 100.0 | 100.0 | |

*Gender*

Similarly, most of the candidates were single i.e., 60% whereas the rest were married accounting to be only 40% of the entire sample as displayed in Table 2.

Table 2. Marital Status

**Marital Status**

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | Married | 40 | 40.0 | 40.0 | 40.0 |
| | Single | 60 | 60.0 | 60.0 | 100.0 |
| | Total | 100 | 100.0 | 100.0 | |

As a Ph.D. is a higher degree, the age bracket starts from 20 years of age. Around 76% of the sample was between the ages of 20 to 30, followed by 20% within 31-40 and only 4 were either 41 or above. Table 3 displays the categorical distribution of the sample by age.

Table 3. Age

**Age at Entry**

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | 20-30 | 76 | 76.0 | 76.0 | 76.0 |
| | 31-40 | 20 | 20.0 | 20.0 | 96.0 |
| | 41 and Above | 4 | 4.0 | 4.0 | 100.0 |
| | Total | 100 | 100.0 | 100.0 | |

Table 4. Ph.D. completion status

**Ph.D. completion**

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | Completed | 50 | 50.0 | 50.0 | 50.0 |
| | Incomplete | 50 | 50.0 | 50.0 | 100.0 |
| | Total | 100 | 100.0 | 100.0 | |

Coming towards the main sample in consideration, the entire sample is divided into Ph.D. candidates who have completed their degree and who have not. Therefore, an equal sample is considered so that an equal representation can be given to each, based on which, 50% of the sample has completed their degree and 50% of the sample has not as displayed in Table 5.

**Normality**

Table 5. Descriptive Statistics

**Descriptive Statistics**

| | N | Minimum | Maximum | Mean | Std. Deviation | Skewness | | Kurtosis | |
|---|---|---|---|---|---|---|---|---|---|
| | Statistic | Statistic | Statistic | Statistic | Statistic | Statistic | Std. Error | Statistic | Std. Error |
| 1 Letter of Recommendation | 100 | 4 | 9 | 6.94 | 1.324 | -.101 | .241 | -.844 | .478 |
| 2 Letter of Recommendation | 100 | 4 | 9 | 7.00 | 1.303 | -.168 | .241 | -.768 | .478 |
| 3 Letter of Recommendation | 100 | 4 | 9 | 7.06 | 1.324 | -.219 | .241 | -.783 | .478 |
| Student Motivation | 100 | 6 | 9 | 7.82 | .957 | -.334 | .241 | -.848 | .478 |
| Emotional Stability | 100 | 4 | 9 | 6.38 | 1.668 | .121 | .241 | -1.211 | .478 |
| Financial Resources | 100 | 3 | 9 | 5.78 | 1.685 | .044 | .241 | -.754 | .478 |
| Interpersonal Skills | 100 | 4 | 9 | 6.58 | 1.365 | -.123 | .241 | -.639 | .478 |
| Hostility | 100 | 1 | 5 | 2.60 | 1.044 | .217 | .241 | -.422 | .478 |
| Selectors Impression of Applicant | 100 | 5 | 9 | 7.08 | 1.203 | -.227 | .241 | -.970 | .478 |
| College GPA | 100 | 2.75 | 3.97 | 3.5130 | .26591 | -.567 | .241 | .278 | .478 |
| Major GPA | 100 | 3.20 | 4.00 | 3.7778 | .19812 | -.811 | .241 | .222 | .478 |
| GRE Specialty | 100 | 520 | 790 | 652.20 | 70.504 | -.041 | .241 | -.889 | .478 |
| GRE Quantitative | 100 | 550 | 787 | 688.49 | 63.906 | -.343 | .241 | -.899 | .478 |
| GRE Verbal | 100 | 470 | 780 | 631.80 | 71.596 | -.051 | .241 | -.652 | .478 |
| Valid N (listwise) | 100 | | | | | | | | |

Since the Skewness and Kurtosis of the data are between +/-1, the data is considered normal as displayed above in Table 5.

**Correlation**

Table 6 displays that Ph.D. completion significantly correlates with the GRE score of the quantitative portion i.e., -0.608, along with Letter of recommendation-1, Letter of Recommendation-3, Student motivation with a correlation coefficient of -0.592, –0.683, and -0.567 respectively. All these variables are highly correlated with Ph.D. and have a strong negative relationship. College GPA, Letter of Recommendation-2, and Age at entry are moderately correlated and have a moderate negative relationship with Ph.D. with correlation coefficient –0.433, –0.494, and –0.377. From these correlation values, we can assume that these factors determine the Ph.D. completion of a candidate, however, there is no significant relation reported in results with Impression of the Selector.

From the correlation table, we can also observe that the independent variables are highly correlated with each other and have a strong positive relationship between them. This high correlation between the independent variables leads to a multicollinearity issue. The following are the highly correlated independent variables. College GPA and Major GPA with a correlation coefficient of 0.901. GRE quantitative correlates with Letter of Recommendation-1 and Letter of Recommendation- 3 whose correlation coefficients are 0.699 and 0.646. GRE specialty and GRE Verbal with correlation 0.984, GRE quantitative, and student motivation with correlation 0.591. Letter of Recommendation-1, Letter of Recommendation- 2 with student motivation with correlation coefficients 0.501 and 0.567respectively. Letter of Recommendation-1 and Letter of Recommendation- 3 with correlation 0.520.

Table 6. Correlation

**Correlations**

| | | PhD completion | Gender | College GPA | Major GPA | GRE Specialty | GRE Quantitative | GRE Verbal | 1 Letter of Recommendation | 2 Letter of Recommendation | 3 Letter of Recommendation | Student Motivation | Emotional Stability | Financial Resources | Marital Status | Age at Entry | Interpersonal Skills | Hostility | Selectors Impression of Applicant |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Pearson Correlation | PhD completion | 1.000 | .250 | -.433 | -.222 | -.048 | -.608 | -.160 | -.592 | -.494 | -.683 | -.567 | -.012 | -.083 | -.082 | -.377 | -.309 | .462 | -.167 |
| | Gender | .250 | 1.000 | .011 | .061 | .012 | -.099 | -.010 | -.156 | -.354 | -.034 | -.208 | .029 | -.100 | -.238 | .075 | -.044 | .048 | .089 |
| | College GPA | -.433 | .011 | 1.000 | .901 | .040 | .408 | .095 | .397 | .262 | .496 | .136 | -.098 | .102 | .120 | .149 | .251 | -.332 | .126 |
| | Major GPA | -.222 | .061 | .901 | 1.000 | .086 | .295 | .096 | .257 | .019 | .329 | -.088 | -.092 | .154 | .119 | .065 | .100 | -.176 | .080 |
| | GRE Specialty | -.048 | .012 | .040 | .086 | 1.000 | .276 | .984 | .064 | -.114 | .050 | -.078 | .189 | -.016 | .043 | -.328 | -.036 | -.122 | .084 |
| | GRE Quantitative | -.608 | -.099 | .408 | .295 | .276 | 1.000 | .374 | .699 | .433 | .646 | .591 | .108 | .018 | .032 | .170 | .280 | -.386 | .166 |
| | GRE Verbal | -.160 | -.010 | .095 | .096 | .984 | .374 | 1.000 | .156 | .017 | .157 | .018 | .210 | -.022 | .044 | -.275 | .023 | -.185 | .105 |
| | 1 Letter of Recommendation | -.592 | -.156 | .397 | .257 | .064 | .699 | .156 | 1.000 | .351 | .520 | .501 | .047 | -.051 | -.068 | .196 | .299 | -.295 | .130 |
| | 2 Letter of Recommendation | -.494 | -.354 | .262 | .019 | -.114 | .433 | .017 | .351 | 1.000 | .351 | .567 | .009 | .064 | .157 | .262 | .364 | -.297 | .077 |
| | 3 Letter of Recommendation | -.683 | -.034 | .496 | .329 | .050 | .646 | .157 | .520 | .351 | 1.000 | .455 | .118 | .024 | -.025 | .291 | .305 | -.450 | .010 |
| | Student Motivation | -.567 | -.208 | .136 | -.088 | -.078 | .591 | .018 | .501 | .567 | .455 | 1.000 | .031 | -.112 | .017 | .258 | .297 | -.396 | -.022 |
| | Emotional Stability | -.012 | .029 | -.098 | -.092 | .189 | .108 | .210 | .047 | .009 | .118 | .031 | 1.000 | -.193 | -.280 | -.189 | -.160 | .355 | -.025 |
| | Financial Resources | -.083 | -.100 | .102 | .154 | -.016 | .018 | -.022 | -.051 | .064 | .024 | -.112 | -.193 | 1.000 | -.034 | -.156 | .091 | -.234 | .069 |
| | Marital Status | -.082 | -.238 | .120 | .119 | .043 | .032 | .044 | -.068 | .157 | -.025 | .017 | -.280 | -.034 | 1.000 | -.031 | -.132 | -.157 | .020 |
| | Age at Entry | -.377 | .075 | .149 | .065 | -.328 | .170 | -.275 | .196 | .262 | .291 | .258 | -.189 | -.156 | -.031 | 1.000 | -.114 | -.232 | -.035 |
| | Interpersonal Skills | -.309 | -.044 | .251 | .100 | -.036 | .280 | .023 | .299 | .364 | .305 | .297 | -.160 | .091 | -.132 | -.114 | 1.000 | -.346 | .304 |
| | Hostility | .462 | .048 | -.332 | -.176 | -.122 | -.386 | -.185 | -.295 | -.297 | -.450 | -.396 | .355 | -.234 | -.157 | -.232 | -.346 | 1.000 | -.264 |
| | Selectors Impression of Applicant | -.167 | .089 | .126 | .080 | .084 | .166 | .105 | .130 | .077 | .010 | -.022 | -.025 | .069 | .020 | -.035 | .304 | -.264 | 1.000 |
| Sig. (1-tailed) | PhD completion | | .006 | .000 | .013 | .316 | <.001 | .056 | <.001 | <.001 | <.001 | <.001 | .453 | .204 | .210 | <.001 | <.001 | <.001 | .048 |
| | Gender | .006 | | .456 | .273 | .452 | .163 | .460 | .061 | .000 | .368 | .019 | .387 | .160 | .009 | .228 | .331 | .317 | .189 |
| | College GPA | .000 | .456 | | .000 | .346 | .000 | .173 | .000 | .004 | .000 | .089 | .167 | .157 | .116 | .069 | .006 | .000 | .107 |
| | Major GPA | .013 | .273 | .000 | | .196 | .001 | .170 | .005 | .426 | .000 | .191 | .181 | .063 | .119 | .260 | .162 | .040 | .216 |
| | GRE Specialty | .316 | .452 | .346 | .196 | | .003 | .000 | .263 | .129 | .309 | .221 | .030 | .436 | .335 | .000 | .359 | .113 | .204 |
| | GRE Quantitative | .000 | .163 | .000 | .001 | .003 | | .000 | .000 | .000 | .000 | .000 | .142 | .431 | .375 | .045 | .002 | .000 | .049 |
| | GRE Verbal | .056 | .460 | .173 | .170 | .000 | .000 | | .061 | .432 | .060 | .429 | .018 | .415 | .333 | .003 | .409 | .033 | .149 |
| | 1 Letter of Recommendation | .000 | .061 | .000 | .005 | .263 | .000 | .061 | | .000 | .000 | .000 | .321 | .306 | .250 | .026 | .001 | .001 | .099 |
| | 2 Letter of Recommendation | .000 | .000 | .004 | .426 | .129 | .000 | .432 | .000 | | .000 | .000 | .463 | .262 | .059 | .004 | .000 | .001 | .222 |
| | 3 Letter of Recommendation | .000 | .368 | .000 | .000 | .309 | .000 | .060 | .000 | .000 | | .000 | .122 | .406 | .403 | .002 | .001 | .000 | .462 |
| | Student Motivation | .000 | .019 | .089 | .191 | .221 | .000 | .429 | .000 | .000 | .000 | | .381 | .133 | .433 | .005 | .001 | .000 | .412 |
| | Emotional Stability | .453 | .387 | .167 | .181 | .030 | .142 | .018 | .321 | .463 | .122 | .381 | | .027 | .002 | .030 | .056 | .000 | .401 |
| | Financial Resources | .204 | .160 | .157 | .063 | .436 | .431 | .415 | .306 | .262 | .406 | .133 | .027 | | .368 | .061 | .183 | .010 | .249 |
| | Marital Status | .210 | .009 | .116 | .119 | .335 | .375 | .333 | .250 | .059 | .403 | .433 | .002 | .368 | | .381 | .095 | .059 | .420 |
| | Age at Entry | .000 | .228 | .069 | .260 | .000 | .045 | .003 | .026 | .004 | .002 | .005 | .030 | .061 | .381 | | .129 | .010 | .364 |
| | Interpersonal Skills | .001 | .331 | .006 | .162 | .359 | .002 | .409 | .001 | .000 | .001 | .001 | .056 | .183 | .095 | .129 | | .000 | .001 |
| | Hostility | .000 | .317 | .000 | .040 | .113 | .000 | .033 | .001 | .001 | .000 | .000 | .000 | .010 | .059 | .010 | .000 | | .004 |
| | Selectors Impression of Applicant | .048 | .189 | .107 | .216 | .204 | .049 | .149 | .099 | .222 | .462 | .412 | .401 | .249 | .420 | .364 | .001 | .004 | |
| N | (all rows) | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

## Multiple Linear Regression Model -1

From the Table 7 model summary, we observe that the R-value = 0.850, which is the correlation coefficient of the overall model with the dependent variable Ph.D. R square = 0.722, which is also known as the coefficient of determination, explains that the model is a good fit model as 72.2% of the variance in Ph.D. completion can be predicted from the independent variables. Both the R and R square values are good which from which we can assume our model may be a good fit.

From the ANOVA (Analysis of Variance) table, the significance value of the overall model is P<0.001 which explains that the model is statistically significant as it is less when compared to alpha 0.05.

Table 7. Model 1

**Model Summary[b]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | R Square Change | F Change | df1 | df2 | Sig. F Change |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Change Statistics | | | | |
| 1 | .850[a] | .722 | .665 | .291 | .722 | 12.552 | 17 | 82 | <.001 |

a. Predictors: (Constant), Selectors Impression of Applicant, 3 Letter of Recommendation, Marital Status, Financial Resources, GRE Speciality, Gender, Major GPA, Emotional Stability, Interpersonal Skills, Student Motivation, Age at Entry, 1 Letter of Recommendation, 2 Letter of Recommendation, Hostility, GRE Quantitative, College GPA, GRE Verbal

b. Dependent Variable: PhD completion

**ANOVA[a]**

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 18.060 | 17 | 1.062 | 12.552 | <.001[b] |
| | Residual | 6.940 | 82 | .085 | | |
| | Total | 25.000 | 99 | | | |

a. Dependent Variable: PhD completion

b. Predictors: (Constant), Selectors Impression of Applicant, 3 Letter of Recommendation, Marital Status, Financial Resources, GRE Speciality, Gender, Major GPA, Emotional Stability, Interpersonal Skills, Student Motivation, Age at Entry, 1 Letter of Recommendation, 2 Letter of Recommendation, Hostility, GRE Quantitative, College GPA, GRE Verbal

**Coefficients[a]**

| Model | | Unstandardized Coefficients | | Standardized Coefficients | | | 95.0% Confidence Interval for B | | Correlations | | | Collinearity Statistics | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | t | Sig. | Lower Bound | Upper Bound | Zero-order | Partial | Part | Tolerance | VIF |
| 1 | (Constant) | 4.429 | 1.019 | | 4.345 | <.001 | 2.402 | 6.457 | | | | | |
| | Gender | .203 | .071 | .195 | 2.846 | .006 | .061 | .345 | .250 | .300 | .166 | .723 | 1.384 |
| | College GPA | -.709 | .365 | -.375 | -1.940 | .056 | -1.435 | .018 | -.433 | -.210 | -.113 | .091 | 11.032 |
| | Major GPA | .704 | .466 | .277 | 1.511 | .135 | -.223 | 1.630 | -.222 | .165 | .088 | .100 | 9.954 |
| | GRE Speciality | .007 | .004 | 1.007 | 1.638 | .105 | -.002 | .016 | -.048 | .178 | .095 | .009 | 111.593 |
| | GRE Quantitative | .001 | .001 | .066 | .563 | .575 | -.001 | .002 | -.608 | .062 | .033 | .246 | 4.073 |
| | GRE Verbal | -.008 | .004 | -1.168 | -1.881 | .063 | -.017 | .000 | -.160 | -.203 | -.109 | .009 | 113.796 |
| | 1 Letter of Recommendation | -.056 | .034 | -.148 | -1.637 | .105 | -.124 | .012 | -.592 | -.178 | -.095 | .416 | 2.403 |
| | 2 Letter of Recommendation | .074 | .044 | .191 | 1.667 | .099 | -.014 | .161 | -.494 | .181 | .097 | .259 | 3.865 |
| | 3 Letter of Recommendation | -.107 | .039 | -.283 | -2.780 | .007 | -.184 | -.031 | -.683 | -.293 | -.162 | .326 | 3.063 |
| | Student Motivation | -.126 | .051 | -.241 | -2.466 | .016 | -.228 | -.024 | -.567 | -.263 | -.143 | .355 | 2.818 |
| | Emotional Stability | -.015 | .024 | -.049 | -.621 | .536 | -.062 | .032 | -.012 | -.068 | -.036 | .549 | 1.820 |
| | Financial Resources | -.054 | .021 | -.181 | -2.617 | .011 | -.095 | -.013 | -.083 | -.278 | -.152 | .710 | 1.408 |
| | Marital Status | -.113 | .072 | -.111 | -1.567 | .121 | -.256 | .030 | -.082 | -.170 | -.091 | .679 | 1.473 |
| | Age at Entry | -.287 | .076 | -.305 | -3.795 | <.001 | -.438 | -.137 | -.377 | -.387 | -.221 | .525 | 1.904 |
| | Interpersonal Skills | -.032 | .030 | -.086 | -1.062 | .291 | -.091 | .028 | -.309 | -.117 | -.062 | .517 | 1.936 |
| | Hostility | -.036 | .043 | -.074 | -.833 | .407 | -.121 | .050 | .462 | -.092 | -.048 | .423 | 2.364 |
| | Selectors Impression of Applicant | -.051 | .028 | -.121 | -1.777 | .079 | -.107 | .006 | -.167 | -.193 | -.103 | .729 | 1.372 |

a. Dependent Variable: PhD completion

As we have 100 observations, so the degrees of freedom df= number of observations - 1 = 99. With 95% of confidence interval and df=99, p-value =0.05, so the critical t-value is ±1.984. Based on corresponding t-values and p-values we can tell which independent variables are good for our model. This can be done by comparing the critical t-value against the calculated t-value, the calculated t-value should not fall in the interval of critical t-value, also checking the significance level of the independent variables whose p-value<0.05. This helps in interpreting which independent variables to be considered as a good fit for our model.

Upon comparison based on the above discussion, we can understand that the independent variables College GPA, Major GPA, GRE specialty, GRE verbal, GRE quantitative, 1 Letter of recommendation, 2 Letter of recommendation, emotional stability, Marital status, interpersonal skills, Hostility, selectors impression of applicants have p-value more than 0.005 and t-calculated values fall in the interval of ±1.98 which is the t-critical value. Hence, indicating that these variables are not statistically significant and may overfit our model. So, there is a need to drop

these variables and rerun our regression model without considering those variables which violates our assumption of statistical significance.
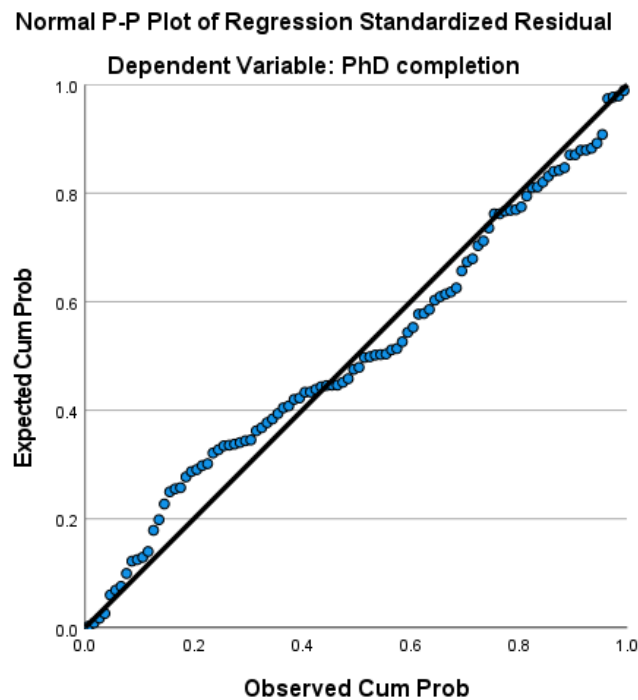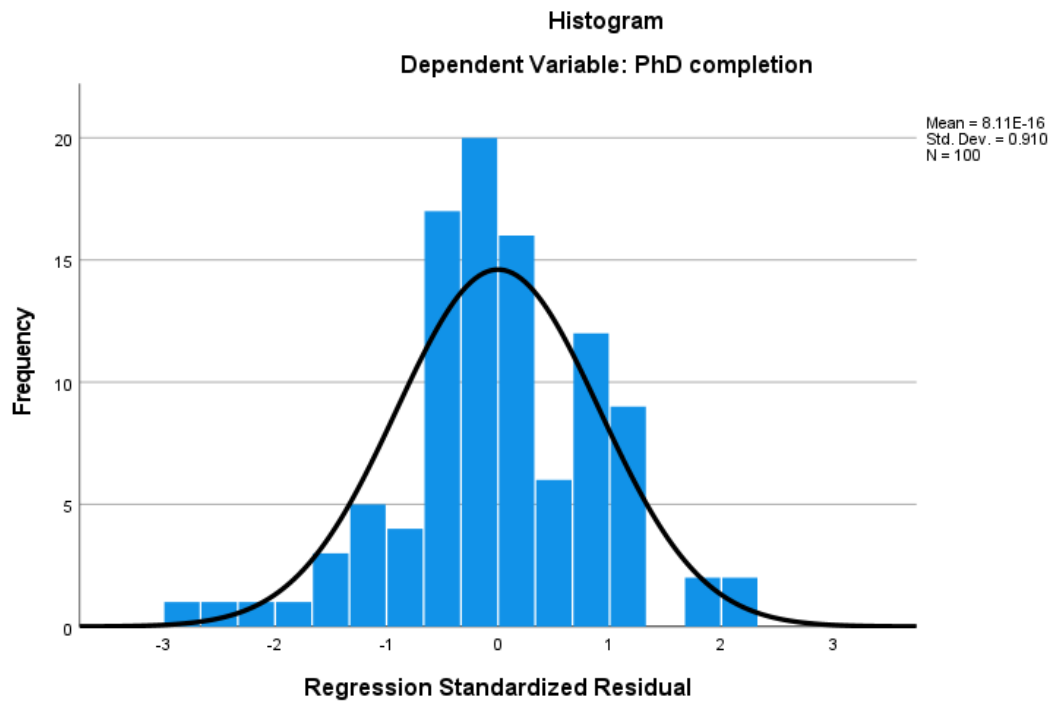
From the histogram and the normal P-P plot, we can infer that the residuals(errors) are not much normally distributed as the residuals do not fall along with the linear line in the normal P-P plot, and the distribution of residuals in the histogram also does not seem normally distributed violating the normally distributed errors assumption.
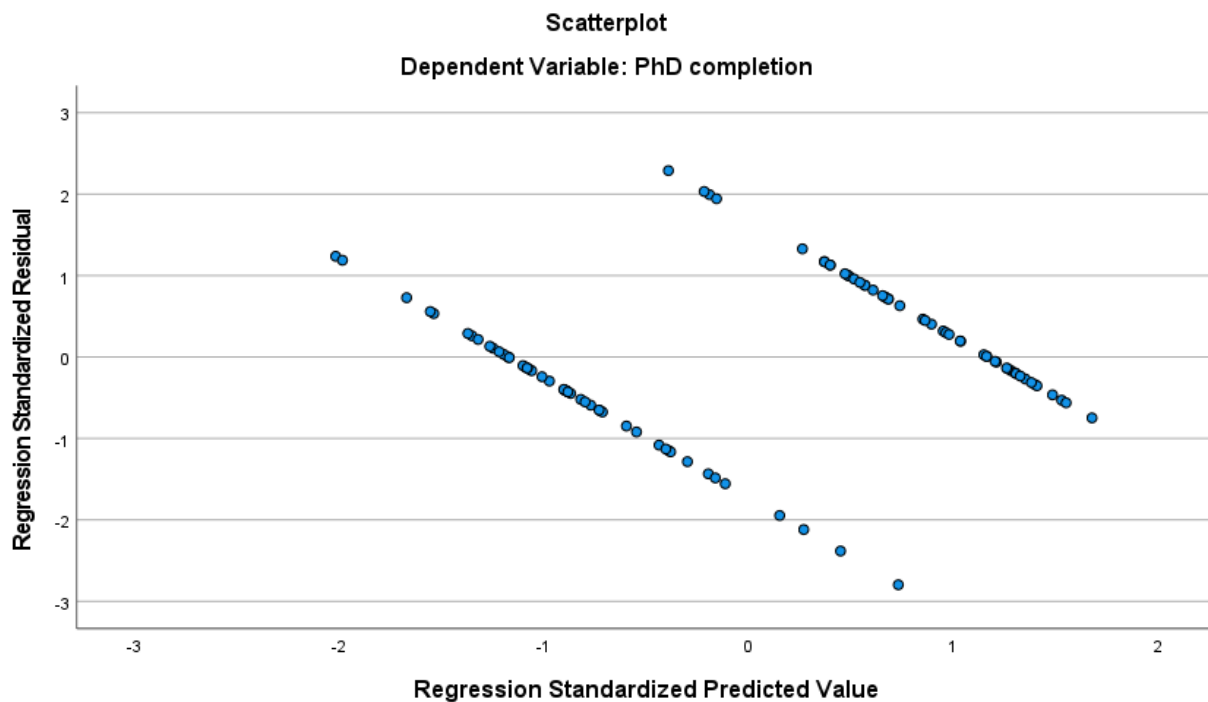


Homoscedasticity, independent errors, Linearity, normally distributed errors Assumptions:

A regression plane shown above tells us about the linearity of the multiple regression with the standardized predictor and residual variables on the x-axis and y-axis respectively which are cantered around Zero (0). We observe a pattern of dots rather than being randomly scattered, showing that the successive residuals are correlated, and the errors are not normally distributed proving that linearity assumption, independent errors, and normally distributed errors assumption are violated. Because of the dots not being normally distributed, it may indicate that the variances of the residuals are not constant which violates the Homoscedasticity assumption.

**Multiple Linear Regression Model 2.**

The following linear regression model is considered by dropping the independent variables from the above regression model-1 based on the p and t-values.

Table 8 Correlations

**Correlations**

| | | PhD completion | Gender | 3 Letter of Recommendation | Student Motivation | Financial Resources | Age at Entry |
|---|---|---|---|---|---|---|---|
| Pearson Correlation | PhD completion | 1.000 | .250 | -.683 | -.567 | -.083 | -.377 |
| | Gender | .250 | 1.000 | -.034 | -.208 | -.100 | .075 |
| | 3 Letter of Recommendation | -.683 | -.034 | 1.000 | .455 | .024 | .291 |
| | Student Motivation | -.567 | -.208 | .455 | 1.000 | -.112 | .258 |
| | Financial Resources | -.083 | -.100 | .024 | -.112 | 1.000 | -.156 |
| | Age at Entry | -.377 | .075 | .291 | .258 | -.156 | 1.000 |
| Sig. (1-tailed) | PhD completion | . | .006 | <.001 | <.001 | .204 | <.001 |
| | Gender | .006 | . | .368 | .019 | .160 | .228 |
| | 3 Letter of Recommendation | .000 | .368 | . | .000 | .406 | .002 |
| | Student Motivation | .000 | .019 | .000 | . | .133 | .005 |
| | Financial Resources | .204 | .160 | .406 | .133 | . | .061 |
| | Age at Entry | .000 | .228 | .002 | .005 | .061 | . |
| N | PhD completion | 100 | 100 | 100 | 100 | 100 | 100 |
| | Gender | 100 | 100 | 100 | 100 | 100 | 100 |
| | 3 Letter of Recommendation | 100 | 100 | 100 | 100 | 100 | 100 |
| | Student Motivation | 100 | 100 | 100 | 100 | 100 | 100 |
| | Financial Resources | 100 | 100 | 100 | 100 | 100 | 100 |
| | Age at Entry | 100 | 100 | 100 | 100 | 100 | 100 |

From the above correlation table, we can infer that the independent variables have a very weak relation among them, indicating that no multicollinearity issue. Letter of recommendation-3 and student motivation are highly correlated with Ph.D., they are exhibiting a strong negative relationship.

**Linear Regression:**

From the Table 8 model summary, we observe that the correlation coefficient R-value = 0.786, coefficient of determination R square = 0.618, when compared with linear regression model-1are low, yet the model is a good fit model as 61.8% of the variance in Ph.D. completion can be predicted from the of the independent variables. There is no change in significance values, hence the overall model is still statistically significant rejecting the null hypothesis with p<0.001 from the ANOVA table.

Table 8 Model 2.

**Model Summary[b]**

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | Change Statistics | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | R Square Change | F Change | df1 | df2 | Sig. F Change |
| 1 | .786[a] | .618 | .598 | .319 | .618 | 30.429 | 5 | 94 | <.001 |

a. Predictors: (Constant), Age at Entry, Gender, Financial Resources, 3 Letter of Recommendation, Student Motivation
b. Dependent Variable: PhD completion

**ANOVA[a]**

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 15.453 | 5 | 3.091 | 30.429 | <.001[b] |
| | Residual | 9.547 | 94 | .102 | | |
| | Total | 25.000 | 99 | | | |

a. Dependent Variable: PhD completion
b. Predictors: (Constant), Age at Entry, Gender, Financial Resources, 3 Letter of Recommendation, Student Motivation
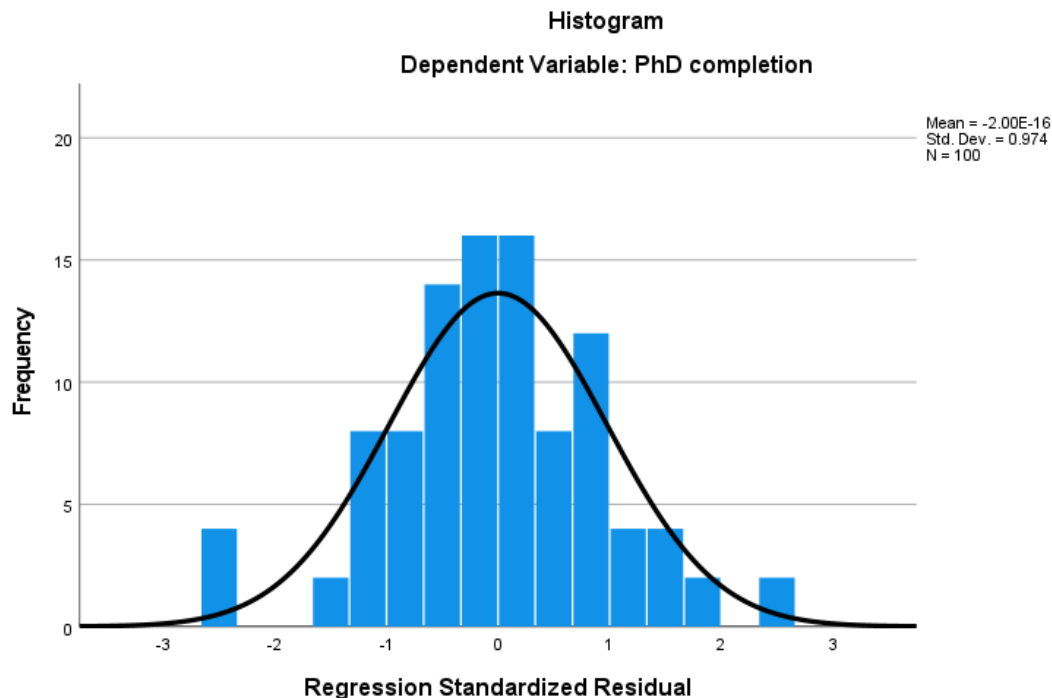
**Coefficients<sup>a</sup>**

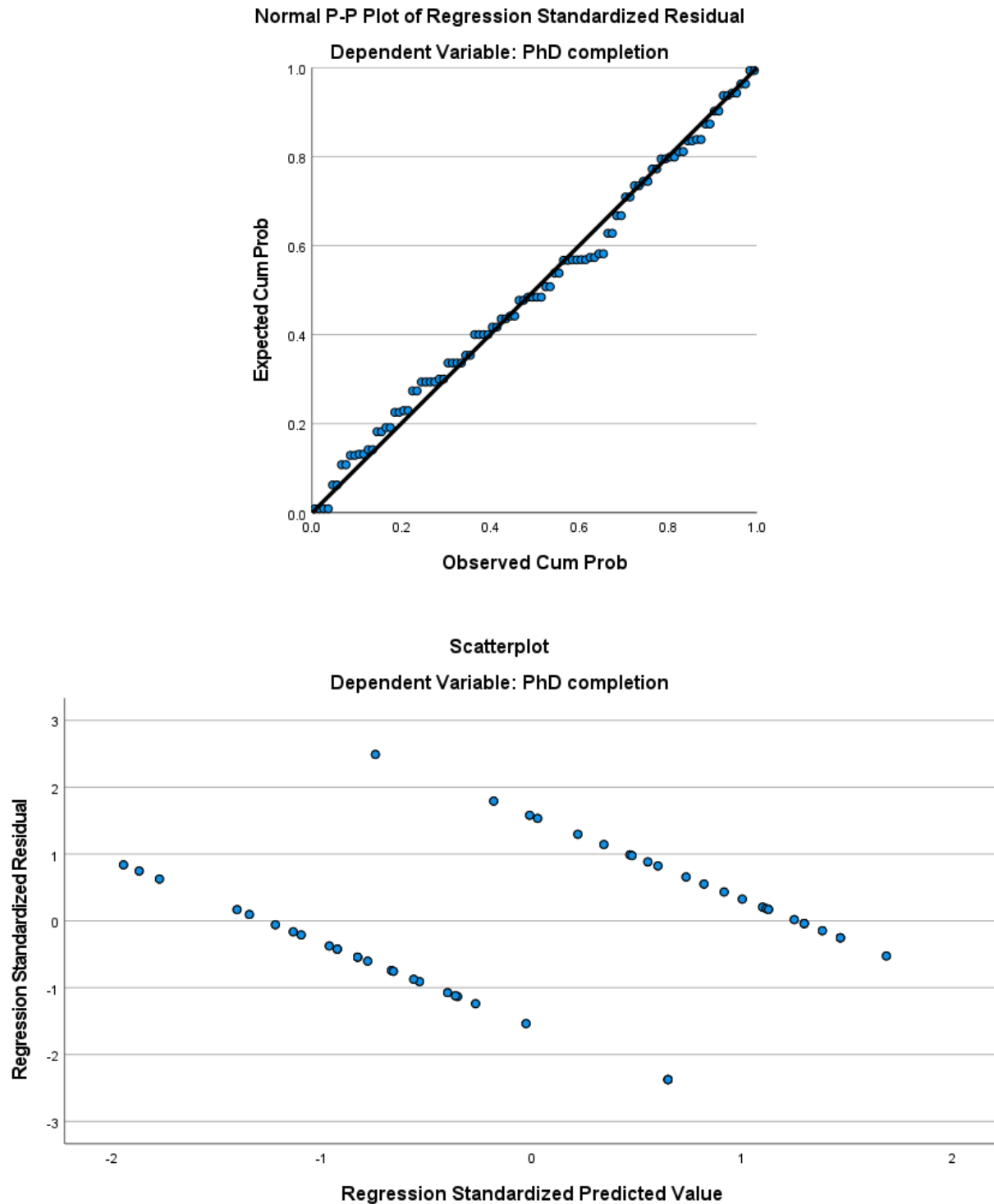| Model | | Unstandardized Coefficients B | Std. Error | Standardized Coefficients Beta | t | Sig. | 95.0% Confidence Interval for B Lower Bound | Upper Bound | Correlations Zero-order | Partial | Part | Collinearity Statistics Tolerance | VIF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | (Constant) | 4.096 | .344 | | 11.916 | <.001 | 3.413 | 4.778 | | | | | |
| | Gender | .189 | .069 | .181 | 2.730 | .008 | .051 | .326 | .250 | .271 | .174 | .924 | 1.082 |
| | 3 Letter of Recommendation | -.188 | .028 | -.496 | -6.739 | <.001 | -.244 | -.133 | -.683 | -.571 | -.430 | .749 | 1.336 |
| | Student Motivation | -.139 | .039 | -.266 | -3.536 | <.001 | -.218 | -.061 | -.567 | -.343 | -.225 | .719 | 1.390 |
| | Financial Resources | -.034 | .020 | -.114 | -1.732 | .086 | -.073 | .005 | -.083 | -.176 | -.110 | .944 | 1.059 |
| | Age at Entry | -.184 | .065 | -.195 | -2.851 | .005 | -.312 | -.056 | -.377 | -.282 | -.182 | .865 | 1.156 |

a. Dependent Variable: PhD completion

The critical t-value remains the same (i.e., ±1.98) which is obtained earlier as the df=99. Applying the same conditions of comparing critical t-value against calculated t-value and corresponding p-values of the independent variables to check which independent variables needed to be dropped.

Depending on that we now conclude that the financial resources independent variable needs to drop. This is because the t-value = -1.732, which falls in the interval of critical t-value, and the corresponding p-value = 0.86 >0.05 indicating that the variable is not statistically significant.

### Histogram
#### Dependent Variable: PhD completion



Mean = -2.00E-16
Std. Dev. = 0.974
N = 100

**Normal P-P Plot of Regression Standardized Residual**

**Dependent Variable: PhD completion**



**Scatterplot**

**Dependent Variable: PhD completion**



From the graphs of the histogram and normal P-P plot between the residuals, we can interpret

that the residuals are normally distributed on the histogram and are aligned along the linear line

on the P-P plot, which is an improvement after dropping the variables. Hence, we may assume that the normally distributed errors assumption is true.

The scatter plot between the standardized predictor values and residuals is an improvement when compared to the model-1 regression plane. Yet this still needs to be improved as it violates the homoscedasticity, linearity, independent errors assumptions because the residuals are not randomly scattered.

## Multiple Linear Regression Model-3

We are now considering a third regression model by dropping the financial resources independent as it violated the assumption.

**Correlation**

**Table 9 Correlations**

**Correlations**

| | | PhD completion | Gender | 3 Letter of Recommendation | Student Motivation | Age at Entry |
|---|---|---|---|---|---|---|
| Pearson Correlation | PhD completion | 1.000 | .250 | -.683 | -.567 | -.377 |
| | Gender | .250 | 1.000 | -.034 | -.208 | .075 |
| | 3 Letter of Recommendation | -.683 | -.034 | 1.000 | .455 | .291 |
| | Student Motivation | -.567 | -.208 | .455 | 1.000 | .258 |
| | Age at Entry | -.377 | .075 | .291 | .258 | 1.000 |
| Sig. (1-tailed) | PhD completion | . | .006 | <.001 | <.001 | <.001 |
| | Gender | .006 | . | .368 | .019 | .228 |
| | 3 Letter of Recommendation | .000 | .368 | . | .000 | .002 |
| | Student Motivation | .000 | .019 | .000 | . | .005 |
| | Age at Entry | .000 | .228 | .002 | .005 | . |
| N | PhD completion | 100 | 100 | 100 | 100 | 100 |
| | Gender | 100 | 100 | 100 | 100 | 100 |
| | 3 Letter of Recommendation | 100 | 100 | 100 | 100 | 100 |
| | Student Motivation | 100 | 100 | 100 | 100 | 100 |
| | Age at Entry | 100 | 100 | 100 | 100 | 100 |

From the above correlation table, we can infer that Letter of recommendation-3 and student motivation are highly correlated with Ph.D. and are exhibiting a strong negative relationship and there is no violation of the multicollinearity assumption.

**Linear Regression Model-3:**

From the Table 9 model summary, we observe that the correlation coefficient R-value = 0.778, coefficient of determination R square = 0.606, explains that the model is a good fit model as 60.6% of the variance in Ph.D. completion can be predicted from the of the independent variables. There is no change in significance values, hence the overall model is still statistically significant rejecting the null hypothesis with p<0.001 from the ANOVA table. The F-value is also good when compared to the previous model-1 and model-2.

**Table 9 Model-3**

Model Summary[b]

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | R Square Change | F Change | df1 | df2 | Sig. F Change |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Change Statistics | | | | |
| 1 | .778[a] | .606 | .589 | .322 | .606 | 36.517 | 4 | 95 | <.001 |

a. Predictors: (Constant), Age at Entry, Gender, 3 Letter of Recommendation, Student Motivation
b. Dependent Variable: PhD completion

ANOVA[a]

| Model | | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regression | 15.148 | 4 | 3.787 | 36.517 | <.001[b] |
| | Residual | 9.852 | 95 | .104 | | |
| | Total | 25.000 | 99 | | | |

a. Dependent Variable: PhD completion
b. Predictors: (Constant), Age at Entry, Gender, 3 Letter of Recommendation, Student Motivation

Coefficients[a]

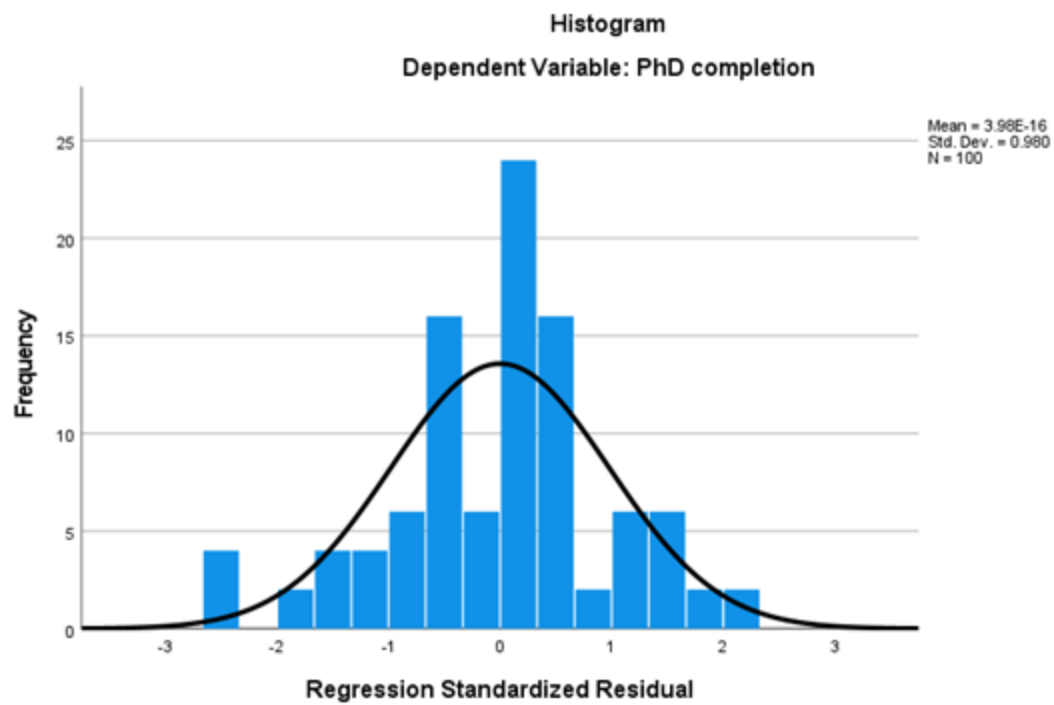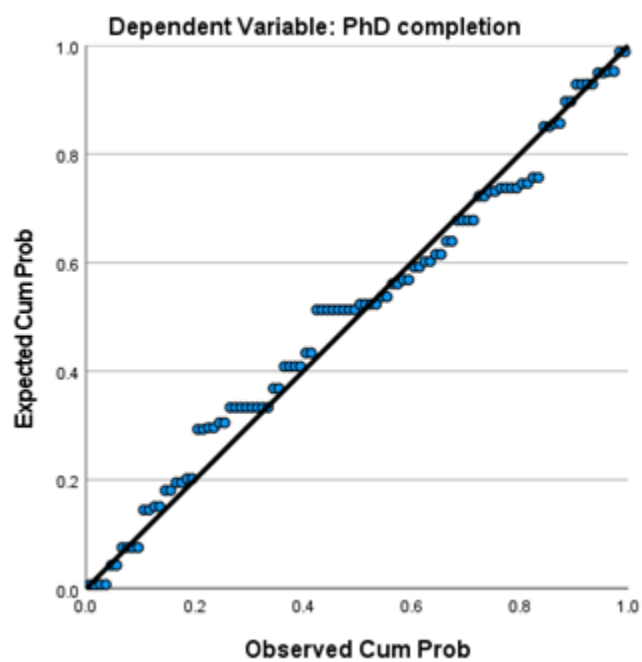| Model | | Unstandardized Coefficients B | Std. Error | Standardized Coefficients Beta | t | Sig. | 95.0% Confidence Interval for B Lower Bound | Upper Bound | Correlations Zero-order | Partial | Part | Collinearity Statistics Tolerance | VIF |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | (Constant) | 3.827 | .310 | | 12.348 | <.001 | 3.212 | 4.442 | | | | | |
| | Gender | .203 | .069 | .195 | 2.923 | .004 | .065 | .340 | .250 | .287 | .188 | .937 | 1.067 |
| | 3 Letter of Recommendation | -.194 | .028 | -.512 | -6.926 | <.001 | -.250 | -.139 | -.683 | -.579 | -.446 | .760 | 1.316 |
| | Student Motivation | -.130 | .039 | -.247 | -3.291 | .001 | -.208 | -.052 | -.567 | -.320 | -.212 | .734 | 1.363 |
| | Age at Entry | -.169 | .065 | -.179 | -2.609 | .011 | -.297 | -.040 | -.377 | -.259 | -.168 | .882 | 1.134 |

a. Dependent Variable: PhD completion

From the coefficients table,

Gender: with the coefficient of 0.203, t-value = 2.923, p=0.04< 0.05, is statistically significant independent variable.Letter of recommendation-3: with the coefficient of –0.194, t-value = -6.926, p=0.001<0.05, is statistically significant independent variable.Student Motivation: with the coefficient of –0.130, t-value = -3.291, p=0.001< 0.05, is statistically significant independent variable.Age at Entry: with the coefficient of –0.169, t-value = -2.609, p=0.011< 0.05, is statistically significant independent variable.Since all the four independent variables are statistically significant, we can now write our regression equation to represent which factors influence Ph.D.
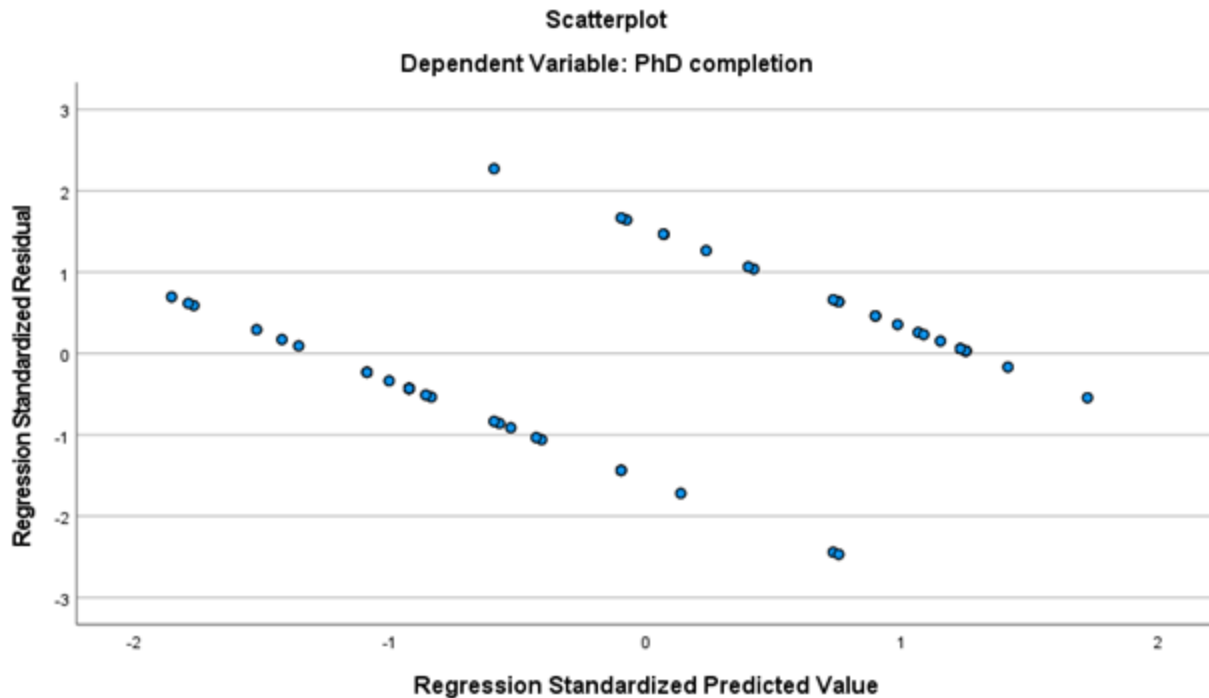
PhD = 3.827 + 2.923*Gender – 0.194*Letter of recommendation -0.130*Student Motivation – 0.169*Age at Entry.

From the regression equation, we can interpret that,Gender: Ph.D. completion is determined by gender with an increase in the factor of 2.923.Letter of Recommendation-3: Ph.D. completion is determined by the letter of recommendation-3 with a decrease in the factor of –0.194. Student Motivation: Ph.D. completion is determined by student motivation with a decrease in the factor of –0.130.Age at entry: Ph.D. completion is determined by age at entry with a decrease in the factor of –0.169.

## Histogram

### Dependent Variable: PhD completion



Mean = 3.98E-16
Std. Dev. = 0.980
N = 100

## Normal P-P Plot of Regression Standardized Residual

### Dependent Variable: PhD completion

**Scatterplot**

**Dependent Variable: PhD completion**



Compared to the Model-2 residual graphs, the model-3 residual plots are far better. From the graphs of the histogram and normal P-P plot between the residuals, we can interpret that the residuals are normally distributed on the histogram and are aligned along the linear line on the P-P plot, which is an improvement after dropping the variables. Hence, we may assume that the normally distributed errors assumption is true.

The scatter plot between the standardized predictor values and residuals is an improvement when compared to the model-2 regression plane. Yet this still needs to be improved because a few dots are still closer to each other. This violates the homoscedasticity, linearity, independent errors assumptions because the residuals are not randomly scattered.

**Discussion**

Based on the results of the regression analysis, Student Motivation, Letter of Recommendation-3, Gender, Age at Entry can easily play the role of a determinant or predictor for Ph.D. completion among candidates. Student Motivation, Letter of Recommendation-3 are strongly correlated with Ph.D. and exhibit a strong negative relationship with correlation coefficients –0.683, -0.567 respectively. Student Motivation can be assumed as the strongest predictor because it is the willpower of an individual himself that makes him capable of achieving greater heights and continue the path towards achieving the Ph.D. degree as well. On the other hand, the letter of recommendations might provide information regarding the caliber and capabilities of the individual, therefore are a great predictor of whether the candidate can complete the Ph.D. degree in time. On the other hand, Age at entry and Gender are moderately correlated with Ph.D., while age shows a negative correlation of –0.377, gender has a positive correlation of 0.250. All the four independent variables Student Motivation, Letter of Recommendation-3, Gender, Age at Entry are statistically significant with each other independently. Among the four-hypothesis derived from the research questions, H2 and H3 hypotheses are satisfied from regression analysis of Model-3.

**Limitations**

In terms of Limitations, only a small sample i.e., a sample was considered to identify whether GPA, GRE scores, and the Interpersonal Abilities of an individual are significant predictors of Ph.D. completion among candidates. Therefore, a small sample cannot be a true representative of the entire country but can be compared to a small locality. Moreover, only 100 candidates have been selected within the locality which is also too short to be a true representative of a large country .

**Recommendations**

The results can further be separated into samples accordingly i.e., people who completed the Ph.D. as well as people who did not complete it. Based on the two different samples, determinants can be predicted based on an independent t-test which would explain statistical variances between the two samples. Moreover, more predictors can also be examined within the study.