FINAL EXAMINATION
SEMESTER 1 SESSION 2017/2018

COURSE CODE : **MANB 1123**

COURSE : **BUSINESS STATISTICS FOR DATA SCIENCE**

PROGRAMME : **MASTER OF SCIENCE BUSINESS INTELLIGENCE AND ANALYTICS (MSc BIA)**

DURATION : **3 HOURS**

DATE :

VENUE :

INSTRUCTIONS TO CANDIDATES:

ANSWER ALL QUESTIONS.

FOR SECTION B, STUDENTS IS ALLOWED TO USE THEIR LAPTOP TO GENERATE THE RESULT. HOWEVER, LAPTOP CAN ONLY BE OPEN AFTER 1 (ONE) HOUR OF THE FINAL EXAM DURATION.

NAME	
I/C NO or PASSPORT NO.	
LECTURER NAME	**DR. NURULHUDA FIRDAUS BT MOHD AZMI**

WARNING !

Students caught copying/cheating during the examination will be liable for disciplinary actions and the Faculty may recommend the student to be expelled from the study.

THIS EXAMINATION PAPER CONSISTS OF 17 PAGES ONLY (INCLUDING COVER PAGE AND APPENDICES).

Plagiarism

(Excerpt from Item 27, UTM Graduate Studies Rules and Regulations)

1) Plagiarism in academic includes producing thesis, dissertation, project report, article, coursework and research findings without acknowledging or referring to the original sources and claiming as one's own work. It involves the submission of the whole or part of the work towards a degree.

2) A Student who committed plagiarism will be penalized based upon decision made by the Faculty Academic Committee.

3) Details of prohibition against plagiarism can be referred to UTM Student Regulations Book.

Examination Misconduct

(Excerpt from Item 28, UTM Graduate Studies Rules and Regulations)

1) A student is suspected of misconduct during examination if:

   a) Giving, receiving or possessing notes or some other materials in various forms relevant to the course during the examination inside and outside of exam hall

   OR

   b) Using the information stated above for the purpose of answering exam questions

   OR

   c) Cheating or attempting to cheat or behaving inappropriately during examination

   OR

   d) Other misconduct set by the University.

2) If the student is found guilty of misconduct by the Faculty Academic Committee and pending the approval of the Senate, the student can be penalized as follow:

   a) Receive a ZERO (0) mark for the examination

   OR

   b) Receive a ZERO (0) mark for the course

   OR

   c) Receive a ZERO (0) mark for all registered courses for the semester

   OR

   d) Suspended from study for a duration specified by the Senate.

SECTION A [50 MARKS]: ANSWER ALL QUESTIONS

QUESTION 1 [5MARKS]

A study was conducted in which 20,211 of 18-year-old male military recruits were given an exam to measure the IQ. In addition, the recruits were asked to disclose their smoking status. An individual was considered a smoker if he smoked at least one cigarette per day. The goal of the study was to determine whether adolescents aged 18 to 21 who smoke have a lower IQ than the non-smokers. It was found that the average IQ of the smokers was 94, while the average IQ of the non-smokers was 101. The researchers concluded that lower IQ individuals are more likely to choose to smoke; not that smoking makes people less intelligent. Identify the following:

(a) The objective of the analysis **[1 Marks]**

(b) The population being studied and the sample **[2 Marks]**

(c) The descriptive statistics of this study **[1 Marks]**

(d) The conclusions of the analysis **[1 Marks]**

QUESTION 2 [5 MARKS]

For each of the following situation, decide the type of **SAMPLING METHOD** that is suitable to use:

a) "National Association of Accountants (NAA) is considered establishing a code of ethics. To determine the opinion of its 20,000 members, a questionnaire was sent to a sample of 500 members." **[1 Mark]**

b) "A farmer divides his orchard into 50 subsections, randomly selects 4, and samples all the trees within the 4 subsections to approximate the yield of his orchard." **[1 Mark]**

c) "A survey of customers entering IKEA shopping mall in Malaysia". **[1 Mark]**

d) "To determine customer opinion of its boarding policy, Air Asia Airlines randomly selects 60 flights during a certain week and surveys all passengers on the flights." **[1 Mark]**

e) "A survey regarding download time on a certain website is administered on the Internet by a market research firm to anyone who would like to take." **[1 Mark]**

QUESTION 3 [7 MARKS]

The PIE CHART shown in **Figure 1** depicts the approaches people use to avoid getting the flu. Answer the following questions:

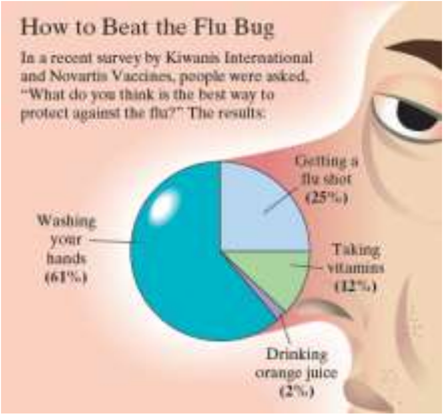


Figure 1: Approaches People Use to Avoid the Flu

a) When do you think pie chart is useful to be used? Explain your answer. [2 Marks]

b) Do you think a single pie chart can be used to compare different sample? Justify your answer. [3 Marks]

c) Identify the **LEAST** common approach people use to avoid the flu? What percentage of the population chooses that method? [1 Marks]

d) What percentage of the population thinks **FLU SHOTS** are the best way to beat the flu? [1 Marks]

QUESTION 4 [8 MARKS]

GetConnected is a new telco company operated in Malaysia since 2 years ago. It has a support center where customers can call to get questions answered about their cell phone account. The manager in charge of the support center has recently conducted a study in which she surveyed 2,300 customers. The customers who called the support center were transferred to a third party who asked the customer series of questions.

a) Choose either the data generated from this study will be considered cross-sectional or time-series? Give your reasons. **[3 Marks]**

b) One of the questions which were asked to the customers was:

"How many minutes they had been on hold waiting to get through to a support person?"

Distinguish the data measurement level (hierarchy) obtained from this question. Give your reasons. **[3 Marks]**

c) Another question which was asked is about the rating of the services on a scale of 1 to 7, with 1 being the worst possible services and 7 being the best possible services. Identify the data measurement level (hierarchy) achieved from this question. Give your reasons. **[2 Marks]**

QUESTION 5 [5 MARKS]

Refer to **Figure 2** and answer the following questions:

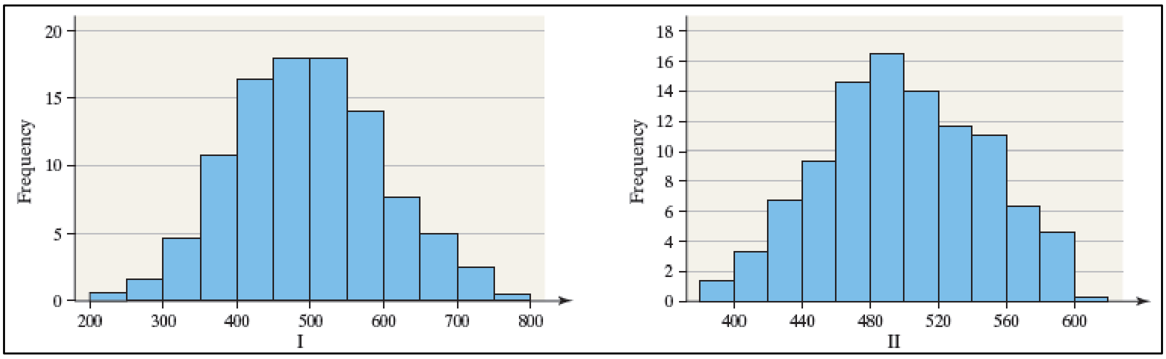


Figure 2: Histogram chart for sample I and sample II

a) Which measure of central tendency would you recommend reporting for the data whose histogram is shown in chart I? Point out your reason? **[2 Marks]**

b) Indicate which sample that has more dispersion? Point out your reason? **[3 Marks]**

QUESTION 6 [12 MARKS]

A researcher had study the prediction model between age difference (husband age minus wife age), y, and literacy rate (percent of the population that is literate), x. The purpose of the model is to predict the relation between the age difference between husband/wives and the percent of a country that is literate. The least-squares regression equation is as below.

$$\hat{y} = -0.0527x + 7.1$$

The prediction model is applied for $18 \leq x \leq 100$.

Source: Xu Zhang and Solomon W. Polachek, State University of New York at Binghamton "The Husband-Wife Age Gap at First Marriage: A Cross-Country Analysis"

Answer the following questions:

a) Interpret the slope. **[1 Marks]**

b) Does it make sense to interpret the y-intercept? Explain. **[3 Marks]**

c) Identify the prediction of age difference between husband/wife in a country where the literacy rate is 25 percent. **[2 Marks]**

d) Would it make sense to use this model to predict the age difference between husband/wife in a country where the literacy rate is 10 percent? Explain **[3 Marks]**

e) The literacy rate in the United States is 99 percent and the age difference between husbands and wives is 2 years. Is this age difference above or below the average age difference among all countries whose literacy rate is 99 percent? **[3 Marks]**

QUESTION 7 [8 MARKS]

Researchers had developed a non-linear regression model to explain the age gap between husbands and wives at first marriage. The model is below:

$$\hat{y} = 3.8483 + 0.0321x_1 + 0.9848x_2 + 0.5391x_3 - 0.000145x_4^2$$

where;

y: Age gap at first marriage (male - female)

x_1: Percent of children aged 10 to 14 involved in child labor

x_2: Indicator variable where 1 is an African country, 0 otherwise

x_3: Percent of the population that is Muslim

x_4^2: Percent of the population that is literate

Source: Xu Zhang and Solomon W. Polachek, State University of New York at Binghamton "The Husband Wife Age Gap at First Marriage: A Cross-Country Analysis"

Answer the following questions:

(a) Use the model to predict the age gap at first marriage for an African country where the percent of children aged 10 to 14 who are involved in child labor is 12, the percent of the population that is Muslim is 30, and the percent of the population that is literate is 75. **[2 Marks]**

(b) Formulate the mean difference in age gap between an African country and a non-African country? **[1 Marks]**

(c) Conclude the coefficient of "percent of children aged 10 to 14 involved in child labor." **[1 Marks]**

(d) The coefficient of determination for this model is 0.593. Conclude this result. **[2 Marks]**

(e) The *P*-value for the test H_0: $b_1 = 0$ versus H_1: $b_1 \neq 0$ is 0.008. What would you conclude about this test? **[2 Marks]**

SECTION B [50 MARKS]: ANSWER ALL QUESTIONS

DATA SET CAN BE DOWNLOADED FROM
https://github.com/hudafirdaus/MANB1123/tree/master/Data_Set
GENERATE THE RESULT USING THE STATISTICAL TOOLS AND ANSWER THE QUESTIONS GIVEN.

QUESTION 1 [5 MARKS]

The owner of the A.J. Fitness Center is interested in estimating the difference in mean years that female members have been with the club compared with male members. He wishes to develop a 95% confidence interval estimate. Sample data are in the file called **AJ FITNESS**. Assuming that the sample data are approximately normal and that the two populations have equal variances, develop and interpret the confidence interval estimate. Discuss the result.

QUESTION 2 [5 MARKS]

Cell phones are becoming an integral part of our daily lives. Commissioned by Motorola, a new behavioral study took researchers to nine (9) cities worldwide from New York to London. Using a combination of personal interviews, field studies, and observation, the study identified a variety of behaviors that demonstrate the dramatic impact cell phones are having on the way people interact. The study found, cell phones give people a newfound personal power, enabling unprecedented mobility and allowing them to conduct their business on the go. Interesting enough, gender differences can be found in cell phone use. Women see their cell phone as a means of expression and social communication, whereas males tend to use it as an interactive toy. A cell phone industry spokesman stated that half of all cell phones in use are registered to females.

a. Formulate the appropriate null and alternative hypotheses for testing the industry claim. **[1 Marks]**
b. Based on a random sample of cell phone owners shown in the data file called **CELL PHONE SURVEY**, test the null hypothesis stated in (a) and interpret the significance test. Use the significance level, $\alpha = 0.05$ **[4 Marks]**

QUESTION 3: [6 MARKS]

A survey by the Pew Internet & American Life Project found that 21% of workers with an e-mail account at work say they are getting more spam than a year ago. Suppose a large multinational company, after implementing a policy to combat spam, asked 198 randomly selected employees with e-mail accounts at work whether they are receiving more spam today than they did a year ago. The results of the survey are in the file **SPAM**. At the 0.025 level of significance, can the company conclude that the percentage of its employees receiving more spam than a year ago is smaller than that found by the Pew study? Discuss your findings.

QUESTION 4: [14 MARKS]

An article in *BusinessWeek* presents a list of the 100 companies perceived as having "hot growth" characteristics. A company's rank on the list is the sum of 0.5 times its rank in return on total capital and 0.25 times its sales and profit-growth ranks. The file entitled **GROWTH** contains sales ($million), sales increase (%), return on capital, market value ($million), and recent stock price of the top 20 ranked companies. Answer the following questions:

a) Identify **TWO** variables that are **MOST** highly correlated with the recent stock price and produce the regression equation to predict the recent stock price as a function of the two variables you choose. **[4 Marks]**

b) Identify the value of coefficient of determination and the adjusted coefficient of determination. What can you conclude based from those value? **[3 Marks]**

c) Identify the variable that is **MOST** correlated with the stock price and test to see if it is a significant predictor of the stock price. Use a significance level of 0.10 and the *p*-value approach. **[3 Marks]**

d) Develop a 95% confidence interval estimate for each of the regression coefficients and interpret each estimate. Comment on whether the interpretation of the intercept is relevant in this situation. **[4 Marks]**

QUESTION 5: [20MARKS]

The athletic director of State University is interested in developing a multiple regression model that might be used to explain the variation in attendance at football games at his school. A sample of 16 games was selected from home games played during the past 10 seasons. Data for the following factors were determined:

y : Game attendance

x_1 : Team win/loss percentage to date

x_2 : Opponent win/loss percentage to date

x_3 : Games played this season

x_4 : Temperature at game time

The data collected are in the file called **FOOTBALL**.

a) Based from the correlation matrix developed on these data, identify whether a multiple regression model will be effectively developed from these data. **[1Marks]**

b) Find the percentage of the total variation in the dependent variable is explained by the independent variables in the model? **[2 Marks]**

c) Distinguish the regression model and identify whether the model is statistically significant. **[3 Marks]**

d) Distinguish if any, of the independent variables is statistically significant? Use a significance level of $\alpha = 0.08$ and the p-value approach to conduct these tests. **[2 Marks]**

e) Estimate the standard deviation of the model error and discuss whether this regression model is acceptable as a means of predicting the football attendance at State University at any given game. **[3 Marks]**

f) Define the term *multicollinearity* and indicate the potential problems that multicollinearity can cause for this model. **[3 Marks]**

g) Indicate what, if any, evidence there is multicollinearity problems with the regression model developed in (c). **[3 Marks]**

h) Construct a 95% confidence interval estimate for each of the regression coefficients and interpret each estimate. Comment on whether the interpretation of the intercept is relevant in this situation. **[3 Marks]**

APPENDIX 1: LIST OF STATISTICAL EQUATIONS

1. Relative Frequency = $\frac{f_i}{n}$ where, f_i = frequency of *i*th value of a discrete variable, *n* = total number of the observation

2. Classes = $2^k \geq n$ rule where, *k* = the number of classes, *n* = total number of data values

3. Class Width (W) = $\frac{Largest\ Value - Smallest\ Value}{Number\ of\ classes}$

4. Sample Proportion, $\bar{p} = \frac{x}{n}$ where,
   x = number of items in the sample with the attribute of interest, *n* = sample size

5. Confidence interval estimate for population proportion (p), $\bar{p} \pm z\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$ where
   $\bar{p}$ = sample proportion, *n* = sample size, *z* = critical value from the standard normal distribution for the desired confidence level

6. Standard error for estimating population proportion $\sigma_{\bar{p}} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$ where
   $\bar{p}$ = sample proportion, *n* = sample size, *z* = critical value from the standard normal distribution for the desired confidence level

7. Sample size for estimating population proportion (p), $n = \frac{z^2\bar{p}(1-\bar{p})}{e^2}$ where
   $\bar{p}$ = sample proportion, *n* = sample size, *e* = desired margin error
   z = critical value from the standard normal distribution for the desired confidence level

8. Margin of error for estimating population proportion (p), $e = z\sqrt{\frac{p(1-p)}{n}}$
   p = population proportion, *n* = sample size, *e* = desired margin error
   z = critical value from the standard normal distribution for the desired confidence level

9. Margin of error for estimating population proportion, $e = z\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$
   $\bar{p}$ = population proportion, *n* = sample size, *e* = desired margin error
   z = critical value from the standard normal distribution for the desired confidence level

10. Confidence interval estimate for mean (μ), standard deviation (σ) known =
    $\overline{x} \pm z\sqrt{\frac{\sigma}{n}}$ where
    σ = population standard deviation, *n* = sample size, $\bar{x}$ = sample mean
    z = critical value from the standard normal distribution for the desired confidence level

11. Margin error for estimating for mean (μ), standard deviation (σ) known, $e = z\sqrt{\frac{\sigma}{n}}$ where, $\sqrt{\frac{\sigma}{n}}$ = standard error of the sample distribution
z = critical value from the standard normal distribution for the desired confidence level

12. Sample size for estimating μ, σ unknown, $n = \frac{z^2\sigma^2}{e^2}$ where
σ = population standard deviation, n = sample size, e = desired margin error
z = critical value from the standard normal distribution for the desired confidence level

13. Sample standard deviation for pilot sample, $s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}}$ where x = data element, n = sample size

14. Confidence interval estimate for μ, σ unknown = $\bar{x} \pm t\sqrt{\frac{s}{n}}$ where
$\bar{x}$ = population mean, s = sample standard deviation, n = sample size
t = critical value from the student's t distribution for the desired confidence level

15. Standard error for sampling distribution with σ unknown, $\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}$

16. Standard error of $(\bar{x}_1 - \bar{x}_2)$ when σ_1 and σ_2 known: $\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ where
σ_1^2 and σ_2^2 = variance for population 1 and population 2 respectively
n_1 and n_2 = sample size for population 1 and population 2 respectively

17. Confidence interval estimate for $(\mu_1 - \mu_2)$ when σ_1 and σ_2 are known with independent samples: $(\bar{x}_1 - \bar{x}_2) \pm z\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ where
n_1 and n_2= sample size for population 1 and population 2 respectively,
z = critical value from the standard normal distribution for the desired confidence level
σ_1^2 and σ_2^2 = variance for population 1 and population 2 respectively

18. Confidence interval estimate for $\mu_1 - \mu_2$ when $\sigma_1$andσ_2 are unknown with independent samples: $(\bar{x}_1 - \bar{x}_2) \pm ts_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ where
n = sample size, t = critical value from the student's t distribution for the desired confidence level with degree of freedom equal to $n_1 + n_2 - 2$

19. Pooled Standard Deviation, $S_p = \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}$ where
n_1 and n_2= sample size for population 1 and population 2 respectively,
s_1^2 and s_2^2 = sample variance for population 1 and population 2 respectively

20. Standard error of $(\bar{x}_1 - \bar{x}_2)$ when σ_1 and σ_2 unknown: $\sigma_{\bar{x}_1-\bar{x}_2} = S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$
where
σ_1^2 and σ_2^2 = variance for population 1 and population 2 respectively
n_1 and n_2 = sample size for population 1 and population 2 respectively

21. Paired difference, $d = x_1 - x_2$ where x_1 and x_2 are values from sample 1 and sample 2 respectively

22. Point estimate for population mean paired difference, $\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n}$ where
d_i= *i*th paired difference value and *n* = number of paired difference

23. Sample standard deviation for paired difference, $s_d = \sqrt{\frac{\sum_{i=1}^{n}(d_i-\bar{d}\,)^2}{n-1}}$ where d_i= *i*th paired difference value and $\bar{d}$= mean paired difference

24. Confidence interval estimate for population mean paired difference, μ_d: $\bar{d} \pm t\frac{s_d}{\sqrt{n}}$
where, $\bar{d}$= mean paired difference, s_d = sample standard deviation of paired difference, n = number of paired difference (sample size), *t* = critical value from the student's *t* distribution for the desired confidence level with degree of freedom equal to $n-1$

25. z-test statistics for hypothesis test for μ, with σ known, $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ where
$\bar{x}$ = Sample mean, μ = hypothesized value for the population mean, σ = population standard deviation, *n*=sample size

26. Critical value sample mean $\bar{x}_\alpha$ for hypothesis test for μ, with σ known, $\bar{x}_\alpha = \mu + z_\alpha\frac{\sigma}{\sqrt{n}}$ where
$\bar{x}$ = Sample mean, μ = hypothesized value for the population mean, σ = population standard deviation, *n*=sample size, z_α = critical value from standard normal distribution

27. t-test statistics for hypothesis test for μ, with σ known, t= $\frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$ with degree of freedom *n-1*, where, $\bar{x}$ = Sample mean, μ = hypothesized value for the population mean, s = sample standard deviation where $s = \frac{\sqrt{(x-\bar{x})^2}}{n-1}$, *n*=sample size

28. z-test statistics for hypothesis test for proportion, $z = \frac{\bar{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$ where
$\bar{p}$ = Sample proportion, p = hypothesized value for the population proportion, *n*=sample size

29. Critical value sample proportion $\bar{p}_\alpha$ for hypothesis test for proportion, $\bar{p}_\alpha = p + z_\alpha\sqrt{\frac{p(1-p)}{n}}$where

$\bar{p}$ = Sample proportion, p = hypothesized value for the population proportion, n=sample size, z_α = critical value from standard normal distribution

30. z- test statistics for $(\mu_1 - \mu_2)$ when σ_1 and σ_2 are known using independent samples: $z = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$ where
$(\bar{x}_1 - \bar{x}_2)$ = sample mean from two population
n_1 and n_2= sample size for population 1 and population 2 respectively,
$(\mu_1 - \mu_2)$= hypothesized difference in the population mean
σ_1^2 and σ_2^2 = variance for population 1 and population 2 respectively

31. t- test statistics for $(\mu_1 - \mu_2)$ when σ_1 and σ_2 are unknown using independent samples: t= $\frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ with degree of freedom $n_1 + n_2 - 2$ where,
n_1 and n_2= sample size for population 1 and population 2 respectively,
$(\mu_1 - \mu_2)$= hypothesized difference in the population mean
S_p= Pooled Standard Deviation

32. t- test statistics for $(\mu_1 - \mu_2)$ when σ_1 and σ_2 are unknown using paired samples: t= $\frac{\bar{d}-\mu_d}{\frac{s_d}{\sqrt{n}}}$ with degree of freedom $n - 1$, where,
$\bar{d}$= mean paired difference, μ_d = hypothesized population mean paired difference
s_d = sample standard deviation for paired difference, n = number of paired difference (sample size),

33. t-value for correlation test, $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$, $df = n - 2$ where, *t* number of standard error *r* is from 0, *r* is sample correlation coefficient and *n* is sample size

34. Regression model diagnosis: prediction range for standard error of the estimate is $\pm \boldsymbol{2s_e}$

APPENDIX 2: STATISTICAL TABLES

Table of the standard normal distribution values ($z \leq 0$)

$-z$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.49601	0.49202	0.48803	0.48405	0.48006	0.47608	0.47210	0.46812	0.46414
0.1	0.46017	0.45621	0.45224	0.44828	0.44433	0.44038	0.43644	0.43251	0.42858	0.42466
0.2	0.42074	0.41683	0.41294	0.40905	0.40517	0.40129	0.39743	0.39358	0.38974	0.38591
0.3	0.38209	0.37828	0.37448	0.37070	0.36693	0.36317	0.35942	0.35569	0.35197	0.34827
0.4	0.34458	0.34090	0.33724	0.33360	0.32997	0.32636	0.32276	0.31918	0.31561	0.31207
0.5	0.30854	0.30503	0.30153	0.29806	0.29460	0.29116	0.28774	0.28434	0.28096	0.27760
0.6	0.27425	0.27093	0.26763	0.26435	0.26109	0.25785	0.25463	0.25143	0.24825	0.24510
0.7	0.24196	0.23885	0.23576	0.23270	0.22965	0.22663	0.22363	0.22065	0.21770	0.21476
0.8	0.21186	0.20897	0.20611	0.20327	0.20045	0.19766	0.19489	0.19215	0.18943	0.18673
0.9	0.18406	0.18141	0.17879	0.17619	0.17361	0.17106	0.16853	0.16602	0.16354	0.16109
1.0	0.15866	0.15625	0.15386	0.15151	0.14917	0.14686	0.14457	0.14231	0.14007	0.13786
1.1	0.13567	0.13350	0.13136	0.12924	0.12714	0.12507	0.12302	0.12100	0.11900	0.11702
1.2	0.11507	0.11314	0.11123	0.10935	0.10749	0.10565	0.10384	0.10204	0.10027	0.09853
1.3	0.09680	0.09510	0.09342	0.09176	0.09012	0.08851	0.08692	0.08534	0.08379	0.08226
1.4	0.08076	0.07927	0.07780	0.07636	0.07493	0.07353	0.07215	0.07078	0.06944	0.06811
1.5	0.06681	0.06552	0.06426	0.06301	0.06178	0.06057	0.05938	0.05821	0.05705	0.05592
1.6	0.05480	0.05370	0.05262	0.05155	0.05050	0.04947	0.04846	0.04746	0.04648	0.04551
1.7	0.04457	0.04363	0.04272	0.04182	0.04093	0.04006	0.03920	0.03836	0.03754	0.03673
1.8	0.03593	0.03515	0.03438	0.03363	0.03288	0.03216	0.03144	0.03074	0.03005	0.02938
1.9	0.02872	0.02807	0.02743	0.02680	0.02619	0.02559	0.02500	0.02442	0.02385	0.02330
2.0	0.02275	0.02222	0.02169	0.02118	0.02068	0.02018	0.01970	0.01923	0.01876	0.01831
2.1	0.01786	0.01743	0.01700	0.01659	0.01618	0.01578	0.01539	0.01500	0.01463	0.01426
2.2	0.01390	0.01355	0.01321	0.01287	0.01255	0.01222	0.01191	0.01160	0.01130	0.01101
2.3	0.01072	0.01044	0.01017	0.00990	0.00964	0.00939	0.00914	0.00889	0.00866	0.00842
2.4	0.00820	0.00798	0.00776	0.00755	0.00734	0.00714	0.00695	0.00676	0.00657	0.00639
2.5	0.00621	0.00604	0.00587	0.00570	0.00554	0.00539	0.00523	0.00509	0.00494	0.00480
2.6	0.00466	0.00453	0.00440	0.00427	0.00415	0.00403	0.00391	0.00379	0.00368	0.00357
2.7	0.00347	0.00336	0.00326	0.00317	0.00307	0.00298	0.00289	0.00280	0.00272	0.00264
2.8	0.00256	0.00248	0.00240	0.00233	0.00226	0.00219	0.00212	0.00205	0.00199	0.00193

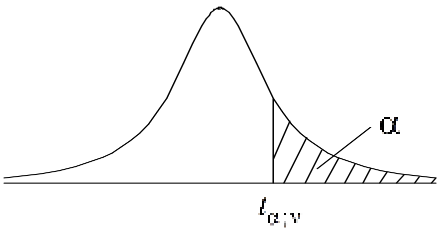
2.9	0.00187	0.00181	0.00175	0.00170	0.00164	0.00159	0.00154	0.00149	0.00144	0.00140
3.0	0.00135	0.00131	0.00126	0.00122	0.00118	0.00114	0.00111	0.00107	0.00104	0.00100
3.1	0.00097	0.00094	0.00090	0.00087	0.00085	0.00082	0.00079	0.00076	0.00074	0.00071
3.2	0.00069	0.00066	0.00064	0.00062	0.00060	0.00058	0.00056	0.00054	0.00052	0.00050
3.3	0.00048	0.00047	0.00045	0.00043	0.00042	0.00040	0.00039	0.00038	0.00036	0.00035
3.4	0.00034	0.00033	0.00031	0.00030	0.00029	0.00028	0.00027	0.00026	0.00025	0.00024
3.5	0.00023	0.00022	0.00022	0.00021	0.00020	0.00019	0.00019	0.00018	0.00017	0.00017

Table of the standard normal distribution values ($z \geq 0$)

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91308	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899

2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983

Table of the Student's *t*-distribution



The table gives the values of $t_{\alpha;\nu}$ where
$\Pr(T_\nu > t_{\alpha;\nu}) = \alpha$, with ν degrees of freedom

α \ ν	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.078	6.314	12.076	31.821	63.657	318.310	636.620
2	1.886	2.920	4.303	6.965	9.925	22.326	31.598
3	1.638	2.353	3.182	4.541	5.841	10.213	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.767
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.291