UNIVERSITY OF LONDON

ST3188

Summer 2020 online assessment guidance

ST3188 Statistical Methods for Market Research

The assessment will be an **open-book take-home online assessment within a 24-hour window**. The requirements for this assessment remain the same as the originally planned closed-book exam, with an expected time/effort of 2 hours.

Candidates should answer the **ONE** question in Section A and **TWO** questions from Section B. Section A carries 40 marks. Questions in Section B carry 30 marks each. Students are strongly advised to divide their time accordingly.

You should complete this assessment using **pen and paper**. Please use **BLACK ink only**.

Handwritten work then needs to be scanned, converted to PDF and then uploaded to the VLE as **ONE individual file** including the coversheet. Each scanned sheet should have your **candidate number** written clearly at the top. Please **do not write your name anywhere** on any sheet.

The paper will be available at **12.00 midday (BST) on Wednesday 29 July 2020**.

You have **until 12.00 midday (BST) on Thursday 30 July 2020** to upload your file into the VLE submission portal. However, you are advised not to leave your submission to the last minute. *We will deduct 5 marks if your submission is up to one hour late, 10 marks if your submission is more than one hour late but less than two hours late (etc.).*

Workings should be submitted for all questions requiring calculations. Any necessary assumptions introduced in answering a question are to be stated.

You may use *any* calculator for any appropriate calculations, but you may not use any computer software to obtain solutions. Credit will only be given if all workings are shown.

If you think there is any information missing or any error in any question, then you should indicate this but proceed to answer the question stating any assumptions you have made.

The assessment has been designed with a duration of 24 hours to provide a more flexible window in which to complete the assessment and to appropriately test the course learning outcomes. As an open-book assessment, the expected amount of effort required to complete all questions and upload your answers during this window is no more than 2 hours. Organise your time well and avoid working all night.

UL20/0542

You are assured that there will be no benefit in you going beyond the expected 2 hours of effort. Your assessment has been carefully designed to help you show what you have learned in the hours allocated.

This is an open book assessment and as such you may have access to additional materials including but not limited to subject guides and any recommended reading. But the work you submit is expected to be 100% your own. Therefore, unless instructed otherwise, you must not collaborate or confer with anyone during the assessment. The University of London will carry out checks to ensure the academic integrity of your work. Many students that break the University of London's assessment regulations did not intend to cheat but did not properly understand the University of London's regulations on referencing and plagiarism. The University of London considers all forms of plagiarism, whether deliberate or otherwise, a very serious matter and can apply severe penalties that might impact on your award. The University of London 2019-20 Procedure for the Consideration of Allegations of Assessment offences is available online at:

https://london.ac.uk/sites/default/files/governance/assessment-offence-procedure-year-2019-2020.pdf

The University of London's Rules for Taking Online Timed Assessments have been included in an update to the University of London General Regulations and are available at:

https://london.ac.uk/sites/default/files/regulations/progregs-general-2019-2020.pdf

SECTION A: Compulsory

1. (a) A travel agency offers customers a range of ways to make holiday bookings – in store, online and through their call centres. Revenues are generated through different commission rates on holidays paid by tour operators (the actual suppliers of the holiday products). The company prides itself on delivering customer satisfaction. Therefore, management is keen to research customer satisfaction levels across the different booking methods and by tour operators (some tour operators offer higher commission margins).

To determine the level of customer satisfaction, the company's management has decided to use a survey of all types of customers and have asked you to devise an appropriate sampling scheme. Explain in detail how each of the following sampling methods could be applied to the overall sampling strategy for this study. Make sure you describe the merits and limitations of each as well as how each would be applied in practice.

    i. Quota sampling

    ii. Snowball sampling

    iii. Systematic sampling

    iv. Stratified sampling.

(20 marks)

(b) Suppose we are interested in estimating the proportion of a population using a simple random sample of size n.

    i. State a suitable estimator of the population proportion as well as its sampling distribution. Mention any assumptions which you make.

    ii. Explain statistically how to determine the minimum sample size necessary to estimate a population proportion to within e units.

    iii. Provide a practical marketing example of a 95% confidence interval for a proportion.

    iv. Explain the purpose of the finite population correction factor (including a formula) and when it should be used.

(20 marks)

SECTION B: Answer two questions. Each question carries equal weight.

2. (a) A retail chain has 5 stores: A, B, C, D and E in a city. The average number of sales transactions (sales volume) based on a random sample has been calculated for three periods during the day (morning, afternoon and evening). The retailer conducts a two-way analysis of variance to investigate whether there appears to be a difference in the average number of sales transactions in different stores and at different times of the day.

   Analyse the selected SPSS output in Figure 1 (spread over the next two pages) and discuss what conclusions can be drawn from the data.

   In your analysis, be sure to address at least the following:

   - Describe the strength of the joint effect of the factors.
   - Test the significance of the variables individually and the interaction between them.
   - Describe the pattern of interaction.

   (20 marks)

   (b) Discuss the reasons for the frequent use of cross-tabulations in market research. What are some of the limitations? Give two examples of cross-tabulations you might expect to find in a market research study.

   (10 marks)

Figure 1

Tests of Between-Subjects Effects

Dependent Variable:Sales volume

Source	Sum of Squares	df	Mean Square	F	Sig.
Time_of_day	1431.667	2	715.833	421.078	.000
Store	1367.467	4	341.867	201.098	.000
Time_of_day * Store	812.333	8	101.542	59.730	.000
Error	25.500	15	1.700		
Total	3636.967	29			

Estimated Marginal Means of Sales volume

Figure 1 (continued)

Estimated Marginal Means of Sales volume

Estimated Marginal Means of Sales volume
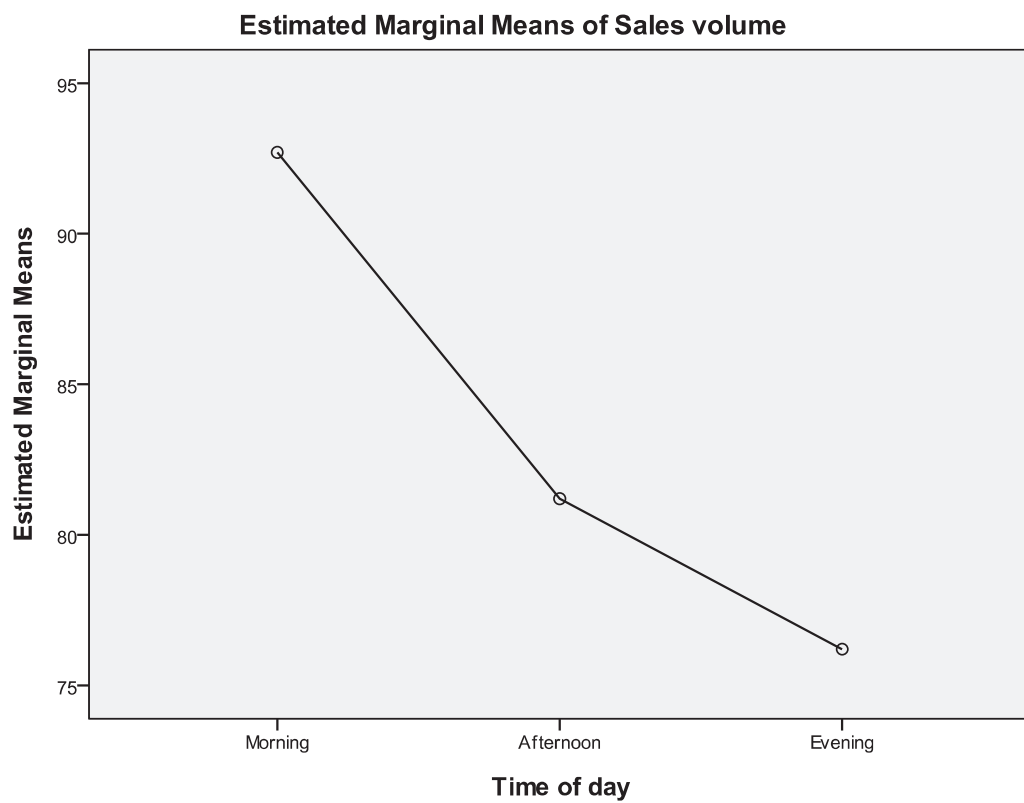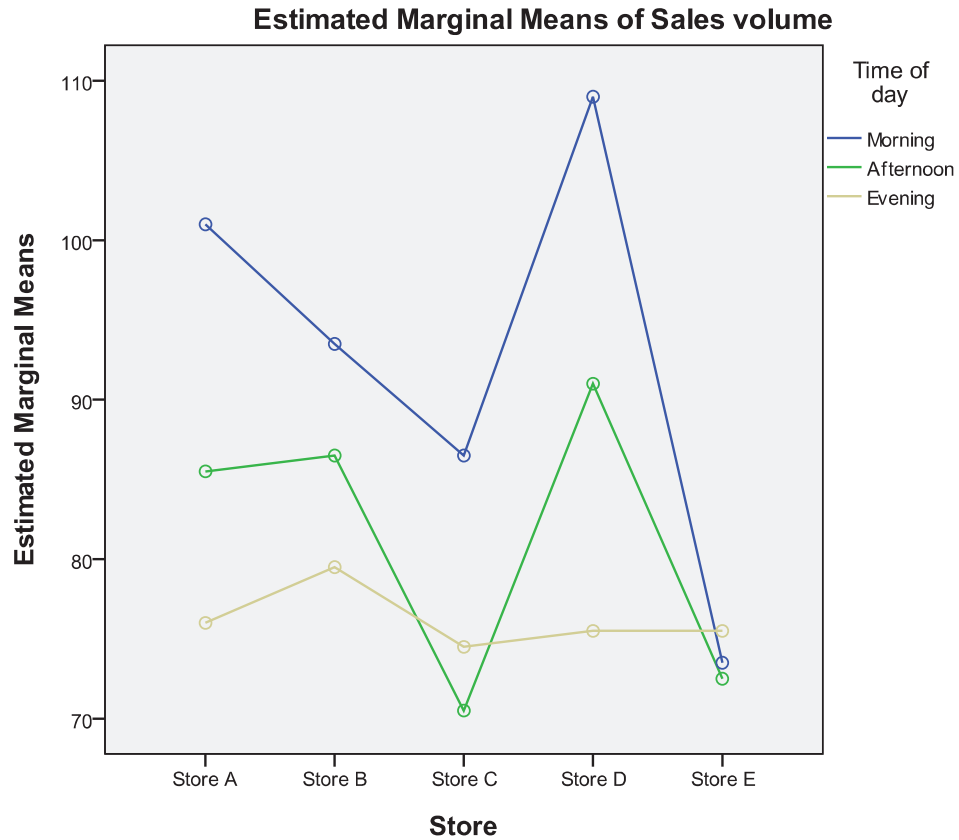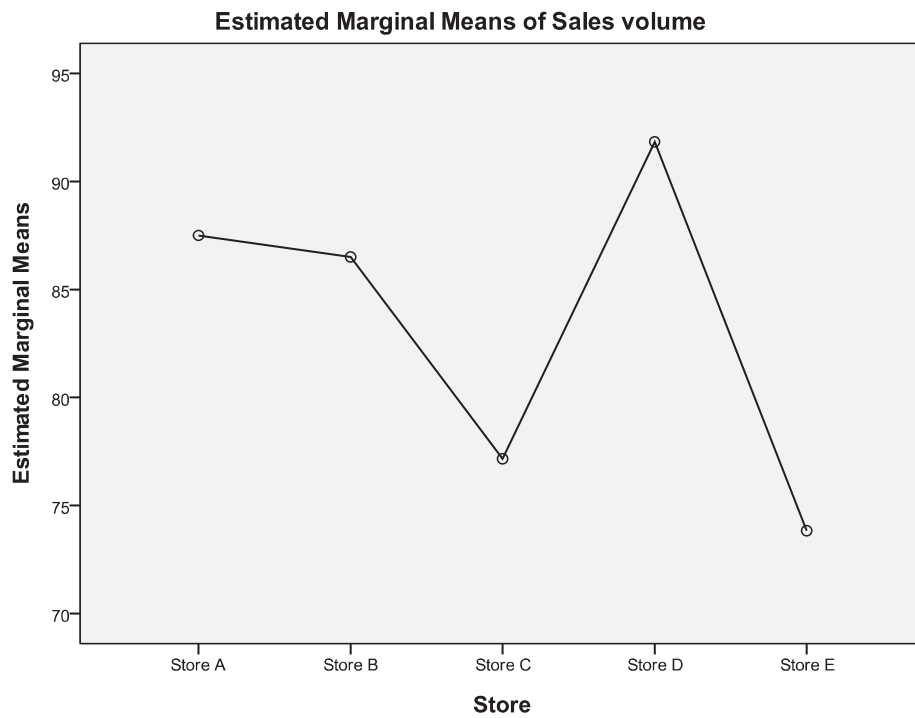
3. (a) A hotel chain decides to use data on existing hotels to determine desirable locations for new hotels. Data on 90 hotels in the chain have been collected in an attempt to understand operating margin. The variables are:

- `Operating margin` – the dependent variable, in %

- `Indoor pool` – whether the hotel has a heated indoor pool, Yes = 1 and No = 0

- `Competitor rooms` – total number of competitor rooms in hotels within 3 miles

- `Distance to competitor` – miles to nearest competitor hotel

- `Office space` – amount of office space available within 3 miles, in thousands of square feet

- `Distance to downtown` – miles to downtown (the central part of a city or town).

Selected SPSS output is provided in Figure 2 (on the next page). Analyse the regression results, making sure you first write out the full regression model, including any assumptions, and the estimated model. In light of the regression results, propose any changes you would make to the model, including any other explanatory variables which you would consider using, or any you would remove from the existing model. Keep in mind the hotel's objective of understanding operating margin to inform its decisions about new hotel sites.

(20 marks)

(b) i. Explain the stepwise regression approach. What is its purpose?

ii. What is multicollinearity? Also, what problems could arise because of the presence of multicollinearity?

**(10 marks)**

Figure 2

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.884[a]	.782	.769	2.9158%

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2557.772	5	511.554	60.169	.000[b]
	Residual	714.164	84	8.502		
	Total	3271.936	89			

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	58.081	2.659		21.840	.000
	Indoor pool	2.316	.671	.184	3.453	.001
	Competitor rooms	-.007	.001	-.536	-10.263	.000
	Distance to competitor	-1.864	.360	-.272	-5.172	.000
	Office space	.019	.002	.512	9.717	.000
	Distance to downtown	.195	.099	.103	1.971	.052

Mean = 2.46E-16
Std. Dev. = 0.972
N = 90

Regression Standardized Residual

UL20/0542
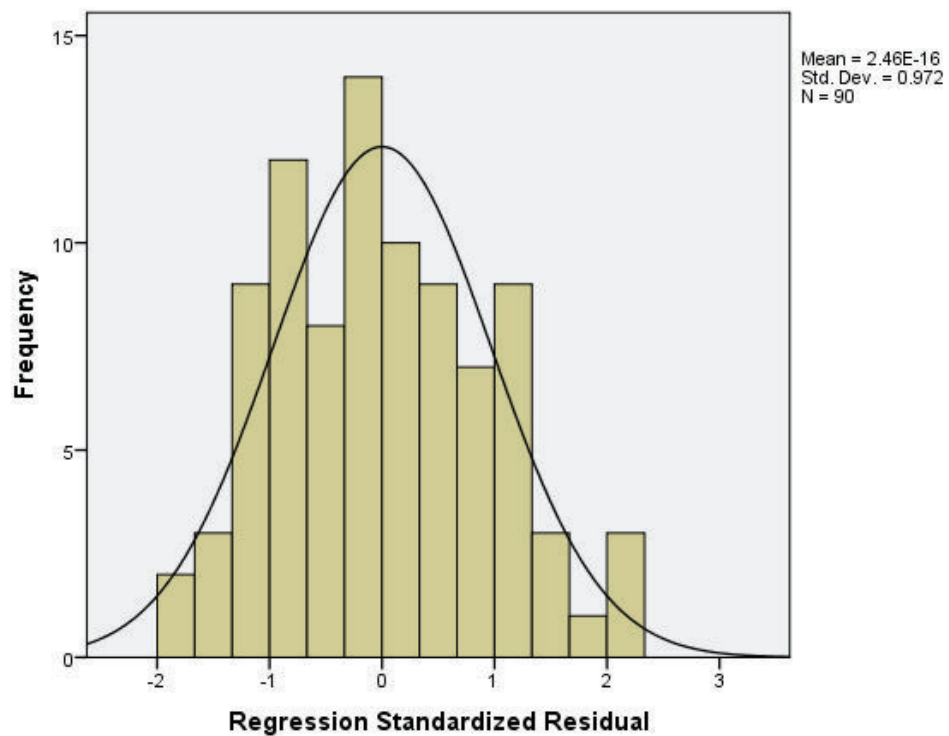
4. (a) Applicants for a position in a firm were asked to score themselves, from 0 to 10, through a questionnaire on the following ten characteristics:

- Ambition

- Appearance

- Drive

- Experience

- Honesty

- Likeability

- Potential

- Salesmanship

- Self-confidence

- Suitability.

When reviewing the questionnaire, because many correlations between the variables are high, it was felt that some of the variables might be confusing, and/or some variables might be redundant. Therefore, a factor analysis was conducted to determine if any underlying factors could be extracted.

Figure 3 (spread over the next two pages) presents selected SPSS output from a factor analysis with principal components extraction, using the varimax rotation procedure. Interpret the output. In your analysis, be sure to address at least the following:

- Explain how you determine the number of factors and interpret the extracted factors.

- Explain qualitatively and quantitatively how the fit of the factor analysis model should be examined.

- Briefly discuss for what modelling purpose(s) any extracted factors could be used.

(20 marks)

(b) How may 'operational data' held by organisations help them to build up an understanding of customer behaviour? Support your answer with examples.

(10 marks)

Figure 3

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.782
Bartlett's Test of Sphericity	Approx. Chi-Square	377.094
	df	45
	Sig.	.000

Communalities

	Initial	Extraction
Appearance	1.000	.432
Likeability	1.000	.841
Self-confidence	1.000	.889
Honesty	1.000	.851
Salesmanship	1.000	.888
Experience	1.000	.848
Drive	1.000	.810
Ambition	1.000	.910
Potential	1.000	.840
Suitability	1.000	.864

Extraction Method: Principal Component Analysis.

Component	Total	Initial Eigenvalues % of Variance	Cumulative %
1	5.336	53.362	53.362
2	1.656	16.562	69.924
3	1.181	11.811	81.735
4	.689	6.891	88.626
5	.336	3.359	91.985
6	.287	2.873	94.858
7	.192	1.925	96.783
8	.132	1.319	98.102
9	.117	1.169	99.271
10	.073	.729	100.000

UL20/0542
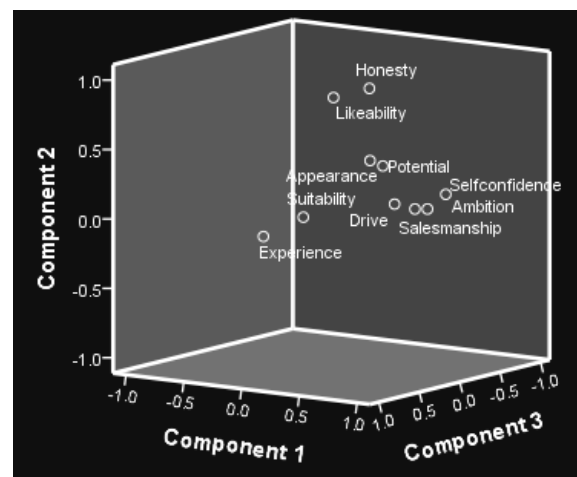
Figure 3 (continued)

Rotated Component Matrix[a]

	Component 1	Component 2	Component 3
Appearance	.454	.443	.171
Likeability	.170	.875	.217
Self-confidence	.914	.206	-.109
Honesty	.178	.879	-.217
Salesmanship	.897	.150	.248
Experience	.055	-.036	.918
Drive	.801	.192	.363
Ambition	.933	.138	.144
Potential	.703	.458	.368
Suitability	.353	.123	.851

Component Score Coefficient Matrix

	Component 1	Component 2	Component 3
Appearance	.040	.180	.023
Likeability	-.184	.529	.108
Self-confidence	.329	-.087	-.233
Honesty	-.101	.525	-.152
Salesmanship	.270	-.115	-.018
Experience	-.141	-.026	.538
Drive	.200	-.064	.069
Ambition	.306	-.133	-.087
Potential	.097	.132	.094
Suitability	-.050	.005	.444

Reproduced Correlations

		Appearance	Likeability	Self-confidence	Honesty	Salesmanship	Experience	Drive	Ambition	Potential	Suitability
Reproduced Correlation	Appearance	.432[a]	.502	.488	.433	.516	.166	.511	.510	.585	.361
	Likeability	.502	.841[a]	.312	.752	.337	.177	.383	.311	.600	.352
	Self-confidence	.488	.312	.889[a]	.367	.823	-.057	.732	.865	.697	.256
	Honesty	.433	.752	.367	.851[a]	.237	-.221	.233	.256	.448	-.013
	Salesmanship	.516	.337	.823	.237	.888[a]	.271	.837	.893	.790	.546
	Experience	.166	.177	-.057	-.221	.271	.848[a]	.370	.178	.360	.797
	Drive	.511	.383	.732	.233	.837	.370	.810[a]	.826	.785	.615
	Ambition	.510	.311	.865	.256	.893	.178	.826	.910[a]	.772	.469
	Potential	.585	.600	.697	.448	.790	.360	.785	.772	.840[a]	.618
	Suitability	.361	.352	.256	-.013	.546	.797	.615	.469	.618	.864[a]

END OF PAPER

UL20/0542