T Level Technical Qualification in Digital Business Services

Employer-set Project (ESP)

Data Technician

Assignment

Assignment brief

v1.1: Specimen assessment materials
September 2021
603/6902/4

Internal reference: DBS-0002-01

T Level Technical Qualification in Digital Business Services Occupational specialism assessment (OSA)

Data Technician

Assignment brief

Assignment

Contents

About this assignment **3**
Introduction 3
Timings 3
Performance outcomes 4
Scenario **5**
About you and your employer 5
About Dynamic Advertising's client 5
The brief **6**
Your role 6
Task 1 ..... **7**
Task 2 ..... **9**
Task 3 ..... **11**
Task 4 ..... **13**
Document information **14**
Change History Record 14

About this assignment

Introduction

This occupational specialism assessment (OSA) is set by NCFE and administered by your provider during a 3 week window. It contains 4 separate tasks which will be completed one after the other during this assessment window.

All 4 tasks will be completed under supervised conditions.

You must complete all tasks in this assignment independently. You are required to sign a declaration of authenticity to confirm that the work is your own. This is to ensure authenticity and to prevent potential malpractice and maladministration. If any evidence was found not to be your own work, it could impact your overall grade.

You will be given a copy of the assignment brief and any relevant supporting information with each task, so you do not have to memorise any information.

Timings

You have a total maximum time of 29 hours to complete all tasks within this assignment, and each task has the following number of hours to complete it:

Task 1 = 5 hours

Task 2 = 10 hours

Task 3 = 8 hours

Task 4 = 6 hours.

Individual tasks must be completed within the timescales stated, but it is up to you how long you spend on each part of the task, therefore manage your time appropriately.

Details on the separate marks available are provided in each task.

You should attempt to complete all of the tasks.

Read the instructions carefully.

Performance outcomes

Marks will be awarded against the skills and knowledge performance outcomes (POs) as follows:

Task 1

This task is divided into 3 parts (part A, B and C) and carries a total of 40 marks.

These are divided between the following performance outcomes:

- PO1: Source, organise and format data securely in a relevant way for analysis (16 marks)
- PO3: Analyse structured and unstructured data to support business outcomes (8 marks)
- PO5: Apply legal, ethical and professional principles when manipulating data (16 marks)

Task 2

This task is divided into 2 parts (part A and B) and carries a total of 52 marks.

These are divided between the following performance outcomes:

- PO1: Source, organise and format data securely in a relevant way for analysis (16 marks)
- PO2: Blend data from multiple sources (20 marks)
- PO3: Analyse structured and unstructured data to support business outcomes (16 marks)

Task 3

This task is divided into 2 parts (part A and B) and carries a total of 40 marks.

These are divided between the following performance outcomes:

- PO3: Analyse structured and unstructured data to support business outcomes (8 marks)
- PO4: Interpret data and communicate a result appropriate to the audience (20 marks)
- PO6: Discover, evaluate and apply reliable sources of knowledge (12 marks)

Task 4

This is a single task and carries a total of 32 marks.

These are divided between the following performance outcomes:

- PO4: Interpret data and communicate a result appropriate to the audience (20 marks)
- PO6: Discover, evaluate and apply reliable sources of knowledge (12 marks)

Scenario

Many businesses use data analytics to plan and organise a marketing strategy. Marketing agencies have access to a large amount of data that can be used to help these businesses plan future activities and strategies.

Work like this within a marketing agency is usually carried out by small teams, each with a team leader who usually has a high level of experience across the agency. Tasks can include analysing existing data provided by the client and bringing this together with publicly available data from social media and demographic sources. This data and initial insights from the team are then sent on to marketing consultants who create a strategy from this information.

About you and your employer

You are a junior data technician in a marketing agency called Dynamic Marketing, which specialises in strategic advice with several high value clients in the clothing and technology sector. You work in a small team of 4 people led by your team leader, Tony Slater.

Jessica McDonald is a corporate manager at Dynamic Marketing and is responsible for monitoring the progress of projects. She does this by having regular meetings with Tony, and occasionally requesting progress reports, which contain basic information and insight to keep her updated. She will get this directly from you, or Tony from time to time.

About Dynamic Advertising's client

Your client is in the technology sector and is the manufacturer and distributor of smart-home IOT appliances, with a focus on the consumer market. The internet of things (IOT) are devices that are embedded with sensors, software, and other technologies for the purpose of connecting and exchanging data with other devices and systems over the internet.

The client's vision statement is: 'To connect the world through innovative and exciting technology'.

Their objectives for the coming year are:

1. Open 2 new smart-home appliance shops. One for their budget range and one for their high-end range.
2. Increase sales of their eight recently launched products.
3. Increase sound system sales.
4. Efficiently stock their security-based products.

The brief

The client has selected Dynamic Marketing to help determine suitable locations to open new stores. They have a range of products which have been proven to appeal to different audiences.

As a data technician, you will provide the client with information to help them succeed in meeting their business objectives. This is done by sourcing and selecting the most appropriate data available, including some research into the current smart-home market and IOT devices, with a focus on devices for the consumer market, along with some selected datasets.

The client is keen to see a proposal before commencing with the project. They are concerned about data sharing, especially as they have offices around the world including Europe, the USA and India. They particularly want to maintain a competitive edge and want to avoid information falling into competitors' hands.

They require insight and a set of recommendations to help them decide their approach, based on research.

The client has told you the following things about their business:

- their products are not very popular with people over 55 years of age
- their sound systems are more popular with under 30s
- customers with larger houses usually buy more products

The client has provided you with a list of their products. Use this list of products, the business objectives and what the client has told you about their business to justify your decisions throughout the project.

Your role

This project requires you to collect and classify data from various sources and of varying types, identify the useful data, and bring together selected data into combined datasets in line with the client's business objectives and their target market for this particular project. Often the original datasets are not structured correctly or contain errors and will require correcting (cleaning) before they can be used by you for the client's purposes.

You must keep in mind all the client's business objectives, even though not all of them will be relevant to every task. This will ensure the work you produce is useful to the client in helping them make these important strategic decisions for the success of their future business.

A large part of your role is to identify patterns and trends in the data supplied, using statistics and logical queries, while checking for errors. Once this is complete, the results can be presented in a summarised dashboard form.

Throughout this process you must keep a log of decisions that you have made, such as when you have had to format different data types, what security measures you had to consider (in line with current relevant legislation), and your chosen methods for verifying and validating your data. The reasons behind your insights and recommendations will be important for the client to help them understand the rationale that sits behind these decisions, which should be data driven.

Task 1

Time limit and marks available

Maximum time allowed = 5 hours (you can use this time how you want during each session, but task 1 must be completed within this time limit).

(40 marks)

Instructions for students

Part A

You are to research the current smart-home market, considering the demographics of the people that buy them, and how they use them, ensuring you use valid sources for this research.

You are required to create a written proposal to meet the client's requirements, which must include tables or charts which show relevant data related to your research on smart home demographics, such as:

- the forecast of the smart-home device market
- the internet of things (IOT) industry, with a focus on devices for the consumer market
- the popularity of smart-home devices by age
- the most popular smart-home devices

Part B

Tony Slater and the client have provided you with some data from various sources. Not all of it will be relevant to the project brief, and some may have errors or need cleaning in order to be useful and reflect the client's requirements.

You are required to:

- select the most appropriate datasets from this selection (which will be given to you by your provider from NCFE)
- discuss, in the form of a written proposal, your choice of datasets and why they are appropriate to the needs of the client and the agency including why you would, or would not use them

Part C

Tony has reminded you to consider relevant laws, regulations and security principles in relation to the client's data. Explain, in a separate section of your proposal which parts of the data are affected by GDPR and the Data Protection Act 2018. You should explain the key principles of data security and also explain what security measures you would put in place when handling this data.

Resources

You will have access to the following resources for all parts of the task, plus the original brief:

- the internet, for research purposes in part A
- task 1 data sets (provided by NCFE)
    - Ages_sctr
    - Ann

  - Client_data_finance
  - Client_data_personal
  - Client_data_sales
  - Client_product_list
  - HomeC
  - Number_of_bedrooms
  - Number_of_rooms
  - Population
  - Raw data
  - Refit_building_survey
  - ukpostcodes
- software applications to select and organise data (Microsoft or Google)
- word processing software (Microsoft or Google)

Evidence required for submission to NCFE

- selected datasets relevant to the project brief
- a single written proposal covering all 3 parts of task 1 (parts A, B and C) which includes the information described in the instructions

Note: you will have access to the internet during this task for the research elements you are required to undertake.

Task 2

Time limit and marks available

Maximum time allowed = 10 hours (you can use this time how you want during each session, but task 2 must be completed within this time limit).

(52 marks)

Instructions for students

The client intends to open two new stores. The client wants one shop to focus on their high-end range for wealthier customers and another to focus on their budget range for less wealthy customers.

The client has decided to locate their high-end store in the KT postcode area and their budget store in the BS postcode area. They are currently undecided which postcode sector to open their respective stores in and want to use a combination of their in-house data and publicly available data to inform their decision.

Tony Slater has provided you with some internal ecommerce sales data and external data sets.

Part A

The client wishes to open their high-end store in a postcode sector where the average house price is over one million, and their budget store in a postcode sector where the average house price is under £250,000.

Tony has asked you to join the external data into one single clean dataset. Make sure the single dataset has appropriate variables which reflect the client's business objectives, as it will eventually help to create a dashboard for the client.

Once cleaned and validated, you must calculate the average house price per postcode sector from the prices dataset. You should exclude postcode sectors without a significant number of sales. The final dataset should also include any calculations which may help you design a dashboard.

Tony would like you to keep a log of your progress and any decisions you make.

This log must include:

- which variables you consider relevant to the business objectives and why
- errors you have found in the datasets
- ways you have validated the data
- which columns you feel are appropriate to the business objectives and why
- the primary keys for each dataset
- data you have removed and why
- the minimum number of sales you considered significant and why
- any calculations and aggregations you have applied to the data
- how you reformatted the data to be joined to the clients' internal data

Include any code or formulas you used to automate the above tasks.

Part B

For this part of the task, the internal data received from the client has been exported from their relational MySQL database. They plan to upload the single dataset you created in part A to their infrastructure. The database will include the following tables:

- CLIENT_PRODUCT_LIST
- CLIENT_DATA_FINANCE

- CLIENT_DATA_PERSONAL
- CLIENT_DATA_SALES
- your new demographics dataset

In addition to part A, Tony has asked you to write a separate additional section in your log, which must include the following:

- describe the normalisation form of this new database, giving a clear explanation of your reasons
- identify the primary, alternate, and foreign keys in each table – write a sentence for each key describing why you have identified it as such
- explain how you reformatted the data to be joined to the external data
- explain how you manipulated date of birth to a format appropriate to the context
- provide a data validation template for each column in your new table which includes **data types** and **constraints**
- explain how you removed any variables from the internal datasets that is not applicable for your analysis

Include any code or formulas you used to automate the above tasks.

Resources

You will have access to the following resources, plus the original brief:

- task 2 data sets (provided by NCFE)
  - Ages_sctr
  - Client_data_finance
  - Client_data_personal
  - Client_data_sales
  - Client_product_list
  - Number_of_bedrooms
  - Number_of_rooms
  - Prices_housetype_key
  - Prices_part_1
  - Prices_part_2
  - Prices_part_3
- software applications to clean and blend data (Microsoft or Google)
- word processing software (Microsoft or Google)

Note: you will not have access to the internet during this task.

Evidence required for submission to NCFE

- single joined data set
- decision log of processes and steps taken as described in the instructions for both parts A and B

Task 3

Time limit and marks available

Maximum time allowed = 8 hours (you can use this time how you want during each session, but task 3 must be completed within this time limit).

(40 marks)

Instructions for students

Part A

Tony has provided you with some internal and external datasets. You have been provided with clean individual datasets and a single joined dataset.

Tony requires you to produce a dashboard for your client. You are not being asked to join the data but may use each dataset how you wish, to help build the dashboard. The dashboard will be presented to the board of directors so must be easy to interact with, be professionally formatted and have obvious variable names and values.

The dashboard should be created using appropriate software and will typically include graphs, tables and filters.

You should automate the dashboard where possible, creating pivot tables and pivot charts which allow the user to interact with the data as they please.

The client wishes to identify trends and patterns in their in-house data and demographic data to help them meet their objectives.

Some of the things the client has told you they are interested in include, but are not limited to:

- high level sales statistics by various customer attributes
  - the popularity of different products by age
  - sales by postcode area and postcode sector
- appropriate demographics of areas they may wish to target – this may include
  - size of the houses
  - the age distribution of a neighbourhood
  - the distribution of property type

You will also need to keep a written decision-making log, similar to task 2, where you must keep a record of the following:

- justification of your choices of the type of visualisation you included
- an explanation of the insights each visualisation provides
- an explanation of how the visualisations are appropriate to the client's objectives

Part B

Your client has also asked if there is a correlation between average house price and percentage of houses with more than 5 bedrooms in a given postcode sector. Perform an appropriate statistical test and explain your findings to the client within the written log. Make sure you reference any appropriate evaluation metric. Explain the reason you chose the statistical test you performed.

Resources

You will have access to the following resources for both parts of the task, plus the original brief:

- task 3 data sets (provided by NCFE)
    - Client_data_personal
    - Client_data_sales
    - Client_product_list
    - HHDemographics
    - Single_set
- software applications to clean and blend data, and create dashboard information
- word processing and spreadsheet software

Evidence required for submission to NCFE

- a dashboard for your client containing detailed information, including trends and patterns you have identified, which help the client towards their business objectives
- a written log containing decisions about selected data to represent, insights and how it relates to the clients' brief
- evidence that an appropriate statistical test has been completed

Task 4

Time limit and marks available

Maximum time allowed = 6 hours (you can use this time how you want within each session, but task 4 must be completed within this time limit).

(32 marks)

Instructions for students

Now that you have completed a comprehensive review of data and researched the smart-home market in line with your client's brief, Tony would like you to draw your conclusions and showcase your results in the form of a recorded presentation internally to Jessica, who can then explain it to the client.

The client is keen to understand any key insights you identify, plus any trends or interesting usage patterns that appear within the data. Your recorded presentation will need an accompanying voice recording so you can explain your findings, insights and recommendations.

You should recommend:

- a postcode sector in which the client should locate each of their new shops, with reasoning
- a plan to increase sound system sales
- how they should stock each shop with their security products

Resources

You will have access to the following resources and equipment, plus the original brief:

- your evidence from task 3 (dashboard, log and statistical test) as a reference to support the creation of your recorded presentation, and help you draw your conclusions
- task 4 data sets (provided by NCFE) – these are the same data sets as task 3 and have been provided so you have all the information to hand in relation to the business context to help you create the most appropriate presentation
- software applications to open a dashboard (Microsoft or Google)
- word processing software (Microsoft or Google)
- presentation software, including equipment for digital voice recording

Evidence required for submission to NCFE

- a screen presentation with voice recording to include:
    - visualisations of completed analysis showing observed trends and patterns via the use of graphs and charts
    - recorded accompanying narrative with explanations, insights and purpose of visualisations, all aligned to the client brief

Document information

The T Level Technical Qualification is a qualification approved and managed by the Institute for Apprenticeships and Technical Education.

Owner: Head of Assessment Design

Change History Record

Version	Description of change	Approval	Date of Issue
v1.0	Post approval, updated for publication.		December 2020
v1.1	Branding and formatting final updates. NCFE rebrand.		September 2021