T-LEVELS | Institute for Apprenticeships & Technical Education

T Level Technical Qualification in Digital Business Services

Occupational specialism assessment (OSA)

Data Technician

Task 2 - Pass

Guide standard exemplification materials

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

Internal reference: DBS-GSEM-06

NCFE

T Level Technical Qualification in Digital Business Services
Occupational specialism assessment

Guide standard exemplification materials

Data Technician

Task 2

Contents

Introduction

The material within this document relates to the Data Technician occupational specialism sample assessment. These exemplification materials are designed to give providers and students an indication of what would be expected for the lowest level of attainment required to achieve a pass or distinction grade.

The examiner commentary is provided to detail the judgements examiners will undertake when examining the student work. This is not intended to replace the information within the qualification specification and providers must refer to this for the content.

In task 2, part A, the student must use the provided datasets and join the data into one single dataset, cleaning the data where required and remembering the client's business objectives. Students must also produce a written decision-making log to keep track of their progress and record any decisions made throughout this task. In part B, students must examine the database and add a section to their decision-making log from part A. This section must include, identify, describe and explain several related points.

After each live assessment series, authentic student evidence will be published with examiner commentary across the range of achievement.

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 2 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

price	date	postcode	property_	new_buil	estate_ty	postcode	postcode	postcode_	Average Hous
£ 242,500	01/05/2019	BS20 7BP	F	N	L	BS20 7	BS20	BS	£ 358,681
£ 215,000	17/05/2019	BS24 7HZ	S	N	F	BS24 7	BS24	BS	£ 220,624
£ 215,000	24/05/2019	BS49 4DR	T	N	F	BS49 4	BS49	BS	£ 319,461
£ 260,000	24/04/2019	BS20 7JY	T	N	L	BS20 7	BS20	BS	£ 358,681
£ 117,500	31/05/2019	BS23 1BA	F	N	L	BS23 1	BS23	BS	£ 180,898
£ 352,500	31/05/2019	BS24 7HL	D	N	F	BS24 7	BS24	BS	£ 220,624
£ 131,000	22/05/2019	BS23 3WH	F	N	L	BS23 3	BS23	BS	£ 179,425
£ 187,000	20/05/2019	BS14 0PJ	S	N	F	BS14 0	BS14	BS	£ 310,272
£ 200,000	31/05/2019	BS41 9EF	F	N	L	BS41 9	BS41	BS	£ 417,518
£ 312,500	14/06/2019	BS31 1WB	T	N	F	BS31 1	BS31	BS	£ 338,601

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

Student evidence

Please see the following files for student evidence for task 2:

- task 2 evidence

Decision making log

Which variables you consider relevant to the business objectives and why

All variables as they might be useful for the dashboard.

Errors you have found in the datasets

There were some pieces of data missing from some of the data sets. This included missing rows or cells of information. If this could not be filled in the row had to be removed. Some information was incorrectly or inconsistently formatted. This had to be made consistent to continue with the processes.

Ways you have validated the data

The power query tool was used to check for correct data types and formatting. If any errors occurred these were fixed, changed, or removed.

Which columns you feel are appropriate to the business objectives and why

Price – because that is what determines the best place for the shops and postcode sector because that is how we find the location.

The super primary keys for each dataset

The super primary key for the prices data set is ID.

	A
	id
94	{8CAC1318-AC35-0253-E053-6B04A8C08E51}
95	{8CAC1318-AC37-0253-E053-6B04A8C08E51}
96	{8CAC1318-AC38-0253-E053-6B04A8C08E51}
97	{8CAC1318-AC3D-0253-E053-6B04A8C08E51}
98	{8CAC1318-AC3F-0253-E053-6B04A8C08E51}
99	{8CAC1318-AC41-0253-E053-6B04A8C08E51}
00	{8CAC1318-AC42-0253-E053-6B04A8C08E51}

Data you have removed and why

No data has been removed as it has been provided by the client and may be important. Only rows with blank values have been removed to make sure the analysis will work.

The minimum number of sales you considered significant and why

All sales have been considered, and they all show income for that specific person.

Any calculations and aggregations you have applied to the data

Mean averages were added for postcode sectors using a pivot table.

How you reformatted the data to be joined to the clients' internal data

All data was formatted in date and currency format. Data was coloured to make it easier to read.

Additional log section (part B)

Describe the normalisation form the relationship schema type of this new database and the normalisation form, giving a clear explanation of your reasons

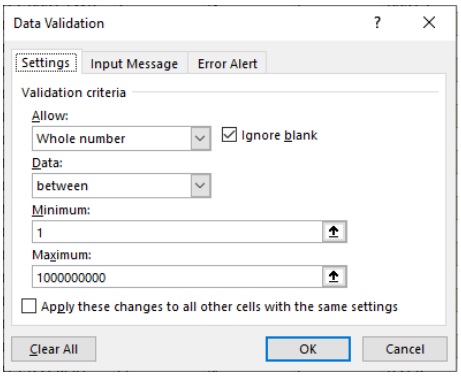
Each table has been organised so that the primary key is the first field.

Duplicate columns have been removed from the tables where they have been found.

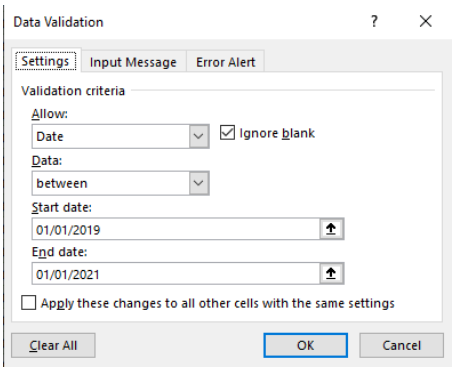
Primary keys and foreign keys have been created to create the relationships between the tables.

Data validation has been used to make sure that the data in each column is correct.

- the data in the price column has been set to be a whole number



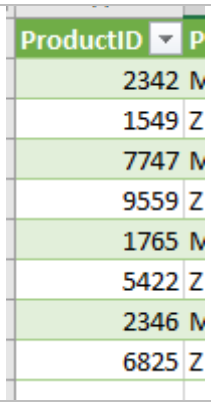
- the data in the date column must be a date

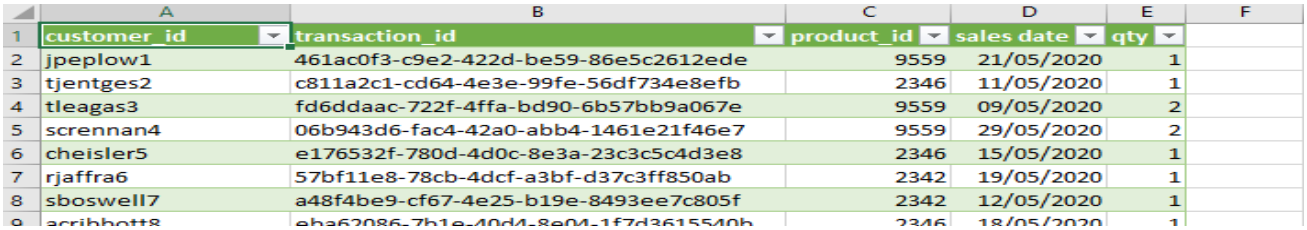


- the data in the property column must be a single character from a specific set
- the data in the new build column must be a single character from a specific set
- the data in the estate type column must be a single character from a specific set
- the data in the postcode sector column must be a 6 character string including 1 space
- data in the postcode district column must be a string of either 3 or 4 characters
- data in the postcode district column must be a 2 character string

Identify the primary, alternate, and foreign keys in each table – write a sentence for each key describing why you have identified it as such

Table	Key	Explanation
CLIENT_DATA_PERSONAL	**Primary key** customer_id	Each value is unique, and the data is consistent on each table.
CLIENT_DATA_FINANCE	**Primary key** customer_id	 Each value is unique, and the data is consistent on each table.

	Alternate key n/a	There is no other column that gives a unique identifier for the data in a row.
CLIENT_PRODUCT_LIST	**Primary key** ProductID	 This allows cross referencing with client data sales.
	Alternate key ProductID	This is unique for every row in the table, but this would not work well with other tables because it is not unique in all tables.
CLIENT_DATA_SALES	**Primary key** customer_id	Each value is unique, and the data is consistent on each table.
	Foreign key product_id	This is a foreign key because there can only be one primary key but even so, each value is unique and links to the other table.
	Foreign key transaction_id	Each value is unique and can identify the row but there can only be one primary key and this value does not allow matches with other tables.
	Alternate key transaction_id	This would also make a good alternate key because it gives a unique identifier to each row in the table. It would not work for this task because it is not present in other tables and can't be used to link them.

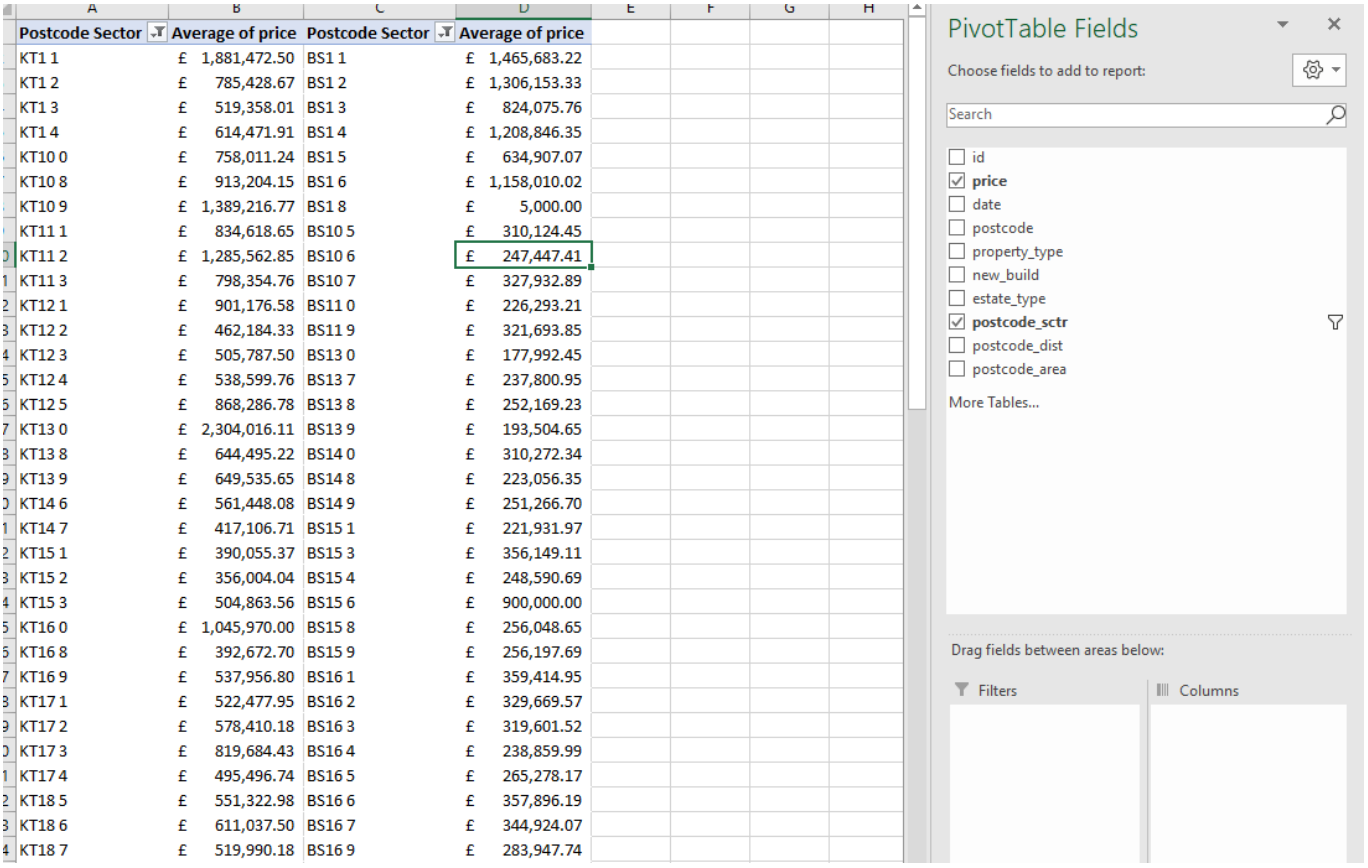


	A	B	C	D	E	F
1	customer_id	transaction_id	product_id	sales date	qty	
2	jpeplow1	461ac0f3-c9e2-422d-be59-86e5c2612ede	9559	21/05/2020	1	
3	tjentges2	c811a2c1-cd64-4e3e-99fe-56df734e8efb	2346	11/05/2020	1	
4	tleagas3	fd6ddaac-722f-4ffa-bd90-6b57bb9a067e	9559	09/05/2020	2	
5	screnan4	06b943d6-fac4-42a0-abb4-1461e21f46e7	9559	29/05/2020	2	
6	cheisler5	e176532f-780d-4d0c-8e3a-23c3c5c4d3e8	2346	15/05/2020	1	
7	rjaffra6	57bf11e8-78cb-4dcf-a3bf-d37c3ff850ab	2342	19/05/2020	1	
8	sboswell7	a48f4be9-cf67-4e25-b19e-8493ee7c805f	2342	12/05/2020	1	
9	acribbett8	eba62086-7b1e-40d4-8e04-1f7d3615540b	2346	18/05/2020	1	

Explain how you reformatted the data to be joined to the external data

Data has been formatted using the tools in excel so that every table has a consistent format, like with the dates. Most of the data could be cleaned using internal features in power query to remove duplicates.

Instead of formulae, I have used a pivot table to work out average house prices in each postcode because this is a much more efficient way to complete this task.

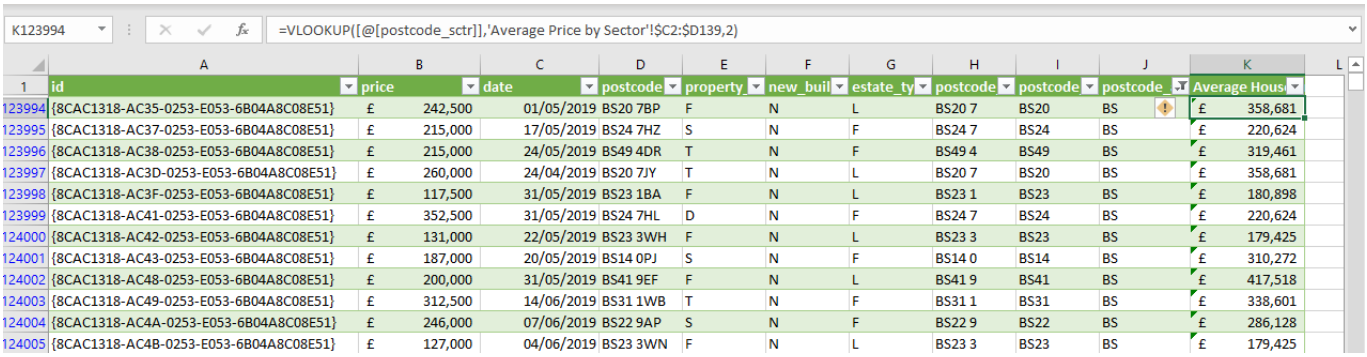


Postcode Sector	Average of price	Postcode Sector	Average of price
KT1 1	£ 1,881,472.50	BS1 1	£ 1,465,683.22
KT1 2	£ 785,428.67	BS1 2	£ 1,306,153.33
KT1 3	£ 519,358.01	BS1 3	£ 824,075.76
KT1 4	£ 614,471.91	BS1 4	£ 1,208,846.35
KT10 0	£ 758,011.24	BS1 5	£ 634,907.07
KT10 8	£ 913,204.15	BS1 6	£ 1,158,010.02
KT10 9	£ 1,389,216.77	BS1 8	£ 5,000.00
KT11 1	£ 834,618.65	BS10 5	£ 310,124.45
KT11 2	£ 1,285,562.85	BS10 6	£ 247,447.41
KT11 3	£ 798,354.76	BS10 7	£ 327,932.89
KT12 1	£ 901,176.58	BS11 0	£ 226,293.21
KT12 2	£ 462,184.33	BS11 9	£ 321,693.85
KT12 3	£ 505,787.50	BS13 0	£ 177,992.45
KT12 4	£ 538,599.76	BS13 7	£ 237,800.95
KT12 5	£ 868,286.78	BS13 8	£ 252,169.23
KT13 0	£ 2,304,016.11	BS13 9	£ 193,504.65
KT13 8	£ 644,495.22	BS14 0	£ 310,272.34
KT13 9	£ 649,535.65	BS14 8	£ 223,056.35
KT14 6	£ 561,448.08	BS14 9	£ 251,266.70
KT14 7	£ 417,106.71	BS15 1	£ 221,931.97
KT15 1	£ 390,055.37	BS15 3	£ 356,149.11
KT15 2	£ 356,004.04	BS15 4	£ 248,590.69
KT15 3	£ 504,863.56	BS15 6	£ 900,000.00
KT16 0	£ 1,045,970.00	BS15 8	£ 256,048.65
KT16 8	£ 392,672.70	BS15 9	£ 256,197.69
KT16 9	£ 537,956.80	BS16 1	£ 359,414.95
KT17 1	£ 522,477.95	BS16 2	£ 329,669.57
KT17 2	£ 578,410.18	BS16 3	£ 319,601.52
KT17 3	£ 819,684.43	BS16 4	£ 238,859.99
KT17 4	£ 495,496.74	BS16 5	£ 265,278.17
KT18 5	£ 551,322.98	BS16 6	£ 357,896.19
KT18 6	£ 611,037.50	BS16 7	£ 344,924.07
KT18 7	£ 519,990.18	BS16 9	£ 283,947.74

I added the average price back into the main sheet by using a vertical lookup.

I used this formula:

=VLOOKUP([@[postcode_sctr]],'Average Price by Sector'!$C$19:$D$165,2)



K123994 =VLOOKUP([@[postcode_sctr]],'Average Price by Sector'!$C2:$D139,2)

	id	price	date	postcode	property	new_buil	estate_ty	postcode	postcode	postcode	Average Hous
123994	{8CAC1318-AC35-0253-E053-6B04A8C08E51}	£ 242,500	01/05/2019	BS20 7BP	F	N	L	BS20 7	BS20	BS	£ 358,681
123995	{8CAC1318-AC37-0253-E053-6B04A8C08E51}	£ 215,000	17/05/2019	BS24 7HZ	S	N	F	BS24 7	BS24	BS	£ 220,624
123996	{8CAC1318-AC38-0253-E053-6B04A8C08E51}	£ 215,000	24/05/2019	BS49 4DR	T	N	F	BS49 4	BS49	BS	£ 319,461
123997	{8CAC1318-AC3D-0253-E053-6B04A8C08E51}	£ 260,000	24/04/2019	BS20 7JY	T	N	L	BS20 7	BS20	BS	£ 358,681
123998	{8CAC1318-AC3F-0253-E053-6B04A8C08E51}	£ 117,500	31/05/2019	BS23 1BA	F	N	L	BS23 1	BS23	BS	£ 180,898
123999	{8CAC1318-AC41-0253-E053-6B04A8C08E51}	£ 352,500	31/05/2019	BS24 7HL	D	N	F	BS24 7	BS24	BS	£ 220,624
124000	{8CAC1318-AC42-0253-E053-6B04A8C08E51}	£ 131,000	22/05/2019	BS23 3WH	F	N	L	BS23 3	BS23	BS	£ 179,425
124001	{8CAC1318-AC43-0253-E053-6B04A8C08E51}	£ 187,000	20/05/2019	BS14 0PJ	S	N	F	BS14 0	BS14	BS	£ 310,272
124002	{8CAC1318-AC48-0253-E053-6B04A8C08E51}	£ 200,000	31/05/2019	BS41 9EF	F	N	L	BS41 9	BS41	BS	£ 417,518
124003	{8CAC1318-AC49-0253-E053-6B04A8C08E51}	£ 312,500	14/06/2019	BS31 1WB	T	N	F	BS31 1	BS31	BS	£ 338,601
124004	{8CAC1318-AC4A-0253-E053-6B04A8C08E51}	£ 246,000	07/06/2019	BS22 9AP	S	N	F	BS22 9	BS22	BS	£ 286,128
124005	{8CAC1318-AC4B-0253-E053-6B04A8C08E51}	£ 127,000	04/06/2019	BS23 3WN	F	N	L	BS23 3	BS23	BS	£ 179,425

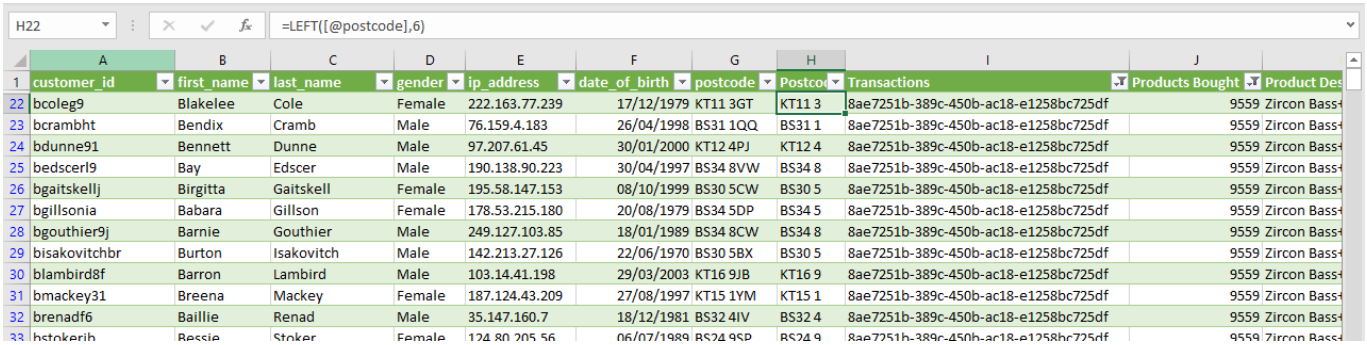
As you can see in the screenshot, there is a formula for the average price which references the pivot table in the previous screenshot, looking up the average price by postcode.

This has the $ sign before the column and row numbers so that they do not change when I fill down the column and they will always look in the right place.

I used a named range in my sheet to make my formula easier to write.

Data types have been set correctly for each column so that numbers, currency and dates are logged appropriately.

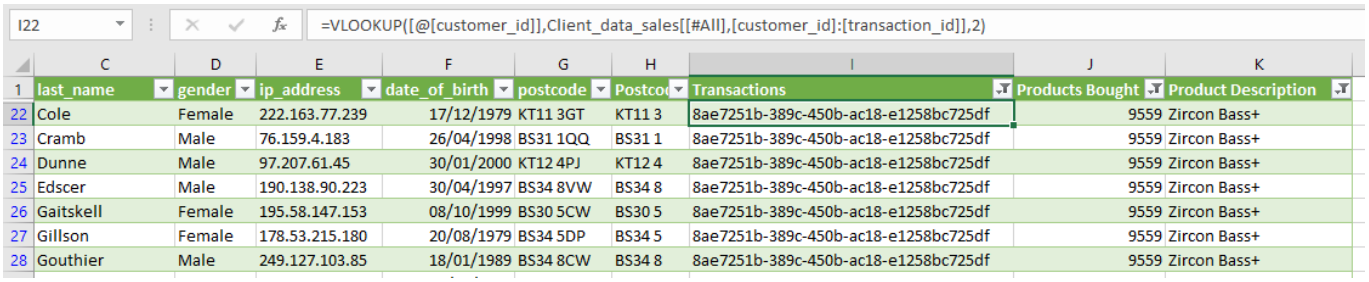
In the client data personal sheet, I added postcode sectors by using the formula: =LEFT([@postcode],6).



	A	B	C	D	E	F	G	H	I	J	
1	customer_id	first_name	last_name	gender	ip_address	date_of_birth	postcode	Postcod	Transactions	Products Bought	Product Des
22	bcoleg9	Blakelee	Cole	Female	222.163.77.239	17/12/1979	KT11 3GT	KT11 3	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass
23	bcrambht	Bendix	Cramb	Male	76.159.4.183	26/04/1998	BS31 1QQ	BS31 1	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass
24	bdunne91	Bennett	Dunne	Male	97.207.61.45	30/01/2000	KT12 4PJ	KT12 4	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass
25	bedscerl9	Bay	Edscer	Male	190.138.90.223	30/04/1997	BS34 8VW	BS34 8	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass
26	bgaitskellj	Birgitta	Gaitskell	Female	195.58.147.153	08/10/1999	BS30 5CW	BS30 5	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass
27	bgillsonia	Babara	Gillson	Female	178.53.215.180	20/08/1979	BS34 5DP	BS34 5	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass
28	bgouthier9j	Barnie	Gouthier	Male	249.127.103.85	18/01/1989	BS34 8CW	BS34 8	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass
29	bisakovitchbr	Burton	Isakovitch	Male	142.213.27.126	22/06/1970	BS30 5BX	BS30 5	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass
30	blambird8f	Barron	Lambird	Male	103.14.41.198	29/03/2003	KT16 9JB	KT16 9	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass
31	bmackey31	Breena	Mackey	Female	187.124.43.209	27/08/1997	KT15 1YM	KT15 1	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass
32	brenadf6	Baillie	Renad	Male	35.147.160.7	18/12/1981	BS32 4IV	BS32 4	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass
33	bstokerjb	Bessie	Stoker	Female	124.80.205.56	06/07/1989	BS24 9SP	BS24 9	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass

This highlighted the first 6 characters of the postcode into a new column and the named range made this easier.

In order to blend the data from the customer data, I have used Vlookup formula using the customer_id as a key to locate other data.



I22 =VLOOKUP([@[customer_id]],Client_data_sales[[#All],[customer_id]:[transaction_id]],2)

	C	D	E	F	G	H	I	J	K
1	last_name	gender	ip_address	date_of_birth	postcode	Postcod	Transactions	Products Bought	Product Description
22	Cole	Female	222.163.77.239	17/12/1979	KT11 3GT	KT11 3	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass+
23	Cramb	Male	76.159.4.183	26/04/1998	BS31 1QQ	BS31 1	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass+
24	Dunne	Male	97.207.61.45	30/01/2000	KT12 4PJ	KT12 4	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass+
25	Edscer	Male	190.138.90.223	30/04/1997	BS34 8VW	BS34 8	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass+
26	Gaitskell	Female	195.58.147.153	08/10/1999	BS30 5CW	BS30 5	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass+
27	Gillson	Female	178.53.215.180	20/08/1979	BS34 5DP	BS34 5	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass+
28	Gouthier	Male	249.127.103.85	18/01/1989	BS34 8CW	BS34 8	8ae7251b-389c-450b-ac18-e1258bc725df	9559	Zircon Bass+

This screenshot shows the Vlookup in action and the data that it generates. The formula is highlighted in the formula bar and the data is showing up in the column. The same method is used in the next 3 columns to do the same thing.

Explain how you manipulated date of birth to a format appropriate to the context

Dates of birth are logged correctly to show date format of year and month. This was done using the functions within excel by automating a query to convert the column to the right format.

A data validation template for each column in your new table which includes data types and constraints

Data validation has been completed as follows with relevant constraints:

Price = currency – only this type of numeric data can be entered.

Date = dates – only this type of numeric data can be entered.

Postcode = string – only text information in the format of a full postcode is allowed.

Property = string – only text is allowed.

New build = string – only text containing the relevant characters is allowed.

Estate type = string – only text containing the relevant characters is allowed.

Postcode area = string – only the text of a relevant length is allowed.

Postcode sector = string – only the text of a relevant length is allowed.

ID = integer – only numerical data is allowed.

Price = integer – only numerical data in a currency format is allowed.

Customer ID = string – only the text of a relevant length is allowed.

Product ID = integer – only a number is allowed in this column.

Sales date = date – only data in the appropriate format of DD/MM/YY is allowed.

Qty = integer – only numbers are allowed.

Data in these columns have only been accepted if they fit these data types. Where data does not meet these data types they have been converted or the rows have been deleted if this is not possible.

Explain how you removed any variables from the internal datasets that is not applicable for your analysis

No data has been removed from the data set in case it is needed and may be useful in the future. Blank and incomplete rows have been removed though. This makes sure they are no unexpected results from the data.

Examiner commentary

Student work is organised logically, reflecting a basic level of application of the skills needed in an employee in this sector.

The student has merged the datasets, showing some understanding of business needs in doing so. Their log shows a small amount of detail and a simple understanding of the procedures needed to complete the task.

The student has demonstrated some knowledge of the techniques needed to complete this task, but this has not been explained in depth. They have used 2 correct formula to manipulate data, given reasons why some tasks have been completed, by using other tools within the software, and have made some use of data transformation. There is a simple explanation of this in their log.

The student has failed to fully remove duplicates and empty rows and has not been able to address the problem of the dataset containing counts of bedroom having more rows than the elements relating to the age profile of the population. Despite the failure to recognise and address this issue, simply appending the values to the end of the sheet, along with other errors and omissions, the student has shown some ability to follow procedures in order to build and test a dataset, combined with some good use of data manipulation though the application of mostly correct formulae and other techniques such as pivot tables. While some errors exist in the average price column where the student has failed to correctly set the cell and column references to absolute references. Despite this, the work shows sufficient skill to place this work at the pass boundary.

Some data cleaning techniques have been applied to the dataset, this has not been thoroughly documented and has not fully addressed issues, such as duplicate values and missing values in the data. The student has demonstrated some routine knowledge and skills that might be needed by workers in this sector, but this is inconsistent.

The student has shown an adequate understanding of keys, giving a basic explanation of these in their decision log. They have explained in basic terms how they have used this knowledge effectively to blend their data sets, by using vlookup formulae. The work meets the standards required of the pass borderline because it shows adequate knowledge of data principles and sufficient data manipulation skills, despite the presence of 1 or 2 errors.

Grade descriptors

The performance outcomes form the basis of the overall grading descriptors for pass and distinction grades.

These grading descriptors have been developed to reflect the appropriate level of demand for students of other level 3 qualifications and the threshold competence requirements of the role and have been validated with employers within the sector to describe achievement appropriate to the role.

Grade	Demonstration of attainment
Pass	The evidence is logical and displays the basic knowledge and skills expected of an employee in this sector in the context of the set brief.
	The student demonstrates theoretical knowledge of the sources, foundations, usage and quality of data that is used for analysis. They are able to carry out routine administrative and analytical tasks using simple datasets.
	The student demonstrates an understanding of data blending techniques and is able to carry out routine data blending tasks.
	The student is able to give a simple explanation of how and why data is analysed by a business. They are able to follow the data process in order to build and test a dataset.
	The student is able to demonstrate understanding of visualisation and communication techniques. They are able to provide evidence of communicating data which is relevant to stated business objectives.
	The student is able to state legal and professional principles that are relevant to the manipulation of data. They are able to carry out routine tasks using data in a way that complies with relevant laws and professional standards.
	The student is able to explain how appropriate sources of information can be selected and evaluated. They are able to search for relevant information and can assess the reliability of the knowledge that they generate.
Distinction	The evidence produced in response to the brief is precise and logical, displaying a secure grasp of the knowledge and skills that would be expected of a new recruit in the industry.
	The student demonstrates a thorough understanding of the sources, foundations, usage and quality of data that is used for analysis. They are able to carry out complex and non-routine administrative and analytical tasks with minimal supervision, using both simple and complex datasets.
	The student demonstrates a secure understanding of a range of data blending techniques and is able to carry out both routine and non-routine data blending tasks competently.
	The student is able to demonstrate a detailed understanding of the reasons why a range of businesses might analyse data. They are able to use their own initiative to follow the data process with minimal supervision in order to build and test a complex dataset in response to a specified business problem.
	The student is able to demonstrate a detailed understanding of a range of visualisation and communication techniques that might be appropriate to a range of organisational needs. They are able to work collaboratively to communicate and visualise data, showing links to business objectives in the materials that they produce.

	The student is able to explain the legal and professional principles that are relevant to a range of different data manipulation tasks. They are able to consistently carry out both routine and non-routine tasks in a way that complies with legal requirements and professional standards.
	The student is able to give a detailed explanation of how to select and evaluate a range of different sources of information for a specific task. They are able to search for data that is appropriate to a given task and can corroborate their findings using appropriate methods to evaluate the suitability of data and making appropriate recommendations for improvements in the collation of data for future tasks.

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	Published final version.		May 2021
v1.1	NCFE rebrand		September 2021