Please check the examination details below before entering your candidate information

Candidate surname	Other names

Pearson Edexcel Level 3 GCE

Centre Number	Candidate Number

Technology and Software

Statistics

Advanced

Topic Test

You must have:
Statistical formulae and tables booklet
Calculator

Total Marks

Candidates may use any calculator allowed by Pearson regulations.
Calculators must not have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B).
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided
  – *there may be more space than you need*.
- You should show sufficient working to make your methods clear.
  Answers without working may not gain full credit.
- Unless otherwise stated, inexact answers should be given to three significant figures.
- Unless otherwise stated, statistical tests should be carried out at the 5% significance level.

Information

- A booklet 'Statistical formulae and tables' is provided.
- There are 9 questions in this question paper. The total mark for this paper is 50.
- The marks for **each** question are shown in brackets
  – *use this as a guide as to how much time to spend on each question*.

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.
- If you change your mind about an answer, cross it out and put your new answer and any working underneath.

1. The table shows estimates of the remaining recoverable oil reserves for the United Kingdom.

	A	B	C	D	E	F	G	H	I	J	K
1		1995	1998	1999	2000	2001	2002	2003	2004	2005	2006
2	Oil (million tonnes)										
3											
4	**Reserves**										
5	Proven	605	685	665	630	605	593	571	533	516	479
6	Probable	765	575	455	380	350	327	286	283	300	298
7	Proven plus Probable	1370	1260	1120	1010	955	920	857	816	816	776
8	Possible	520	540	545	480	475	425	410	512	451	
9	Maximum	1890	1800	1665	1490	1430	1344	1267	1328	1267	1254
10											
11	**Expected level of reserves**										
12	Opening stocks	1975	1675	1535	1370	1235	1160	1192	1180	1212	1162
13	Extraction	-130	-132	-137	-126	-117	-117	-106	-95	-85	-77
14	Other volume changes	-95	-8	-28	-9	42	149	94	127	35	130
15	Closing stocks	1750	1535	1370	1235	1160	1192	1180	1212	1162	1215
16											
17	**Life expectancy (years)**	13	12	10	10	10	10	11	13	14	

[Source: Annual Abstract of Statistics, Office for National Statistics, 2008]

Explain how you would use the spreadsheet to calculate

(a) the value in cell K8,

(1)

(b) the mean level of possible reserves per year.

(2)

(Total 3 marks)

2. The following is a printout of a spreadsheet containing data about the amount of fuel used in electricity generation. The data is copied from Department of Energy and Climate Change.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													Million tonnes of oil equivalent
2			**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**2020**
3	**All generating companies: total fuels**	**KGQQ**	79.41	76.52	77.24	74.52	68.48	66.72	63.67	61.43	59.88	56.14	52.24
4	**Coal**	**KGQR**	25.56	26.03	34.33	31.33	24.01	18.34	7.53	5.55	4.24	1.85	1.47
5	**Oil**	**KGQS**	1.18	0.78	0.73	0.59	0.55	0.61	0.58	0.54	0.49	0.39	0.34
6	**Gas**	**KGQT**	32.43	26.58	18.62	17.70	18.73	18.28	25.63	24.27	23.51	23.45	19.99
7	**Nuclear**	**KGQU**	13.93	15.63	15.21	15.44	13.85	15.48	15.41	15.12	14.06	12.09	10.71
8	**Hydro (natural flow)**	**KGQV**	0.31	0.49	0.46	0.40	0.51	0.54	0.46	0.51	0.47	0.51	0.59
9	**Wind and Solar**	**-**	0.89	1.39	1.82	2.61	3.10	4.11	4.09	5.25	5.98	6.56	7.58
10	**Other fuels**	**-**	5.11	5.62	6.07	6.45	7.73	9.36	9.96	10.18	11.13	11.30	11.56

[Source: www.gov.uk/environment/energy-and-climate-change-evidence-and-analysis]

This spreadsheet is generated from a database with the following schema

Variable	Name	Type	Width	Decimal	Label
1	Fuel_Type	String	8	0	Fuel used in generating electricity
2	Code	String	8	0	Fuel Code
3	Year	Numeric	8	2	Year
4	Fuel_Consumption	Numeric	8	2	Amount of fuel used in electricity generation (Million tonnes of oil equivalent)

(a) Explain how the spreadsheet would be able to calculate the average amount of gas used in electricity generation between 2010 and 2020.

(2)

(b) Explain how you would use the database to obtain a list of fuels in alphabetical order.

(2)

(c) Give one advantage and one disadvantage of using spreadsheet software instead of database software.

(2)

(Total 6 marks)

3. In a campaign to increase the level of exercise to improve child mental health, Bronwyn, a children's health worker, wants to collect information about the ways pupils travel to school.

In her pilot study she wants a sample of 50 pupils from a school in the town where she works. She is granted permission to access the school database for the purposes of her study.

SchoolDatabase

ID	PupilSurname	PupilForename	PupilGender	SchoolYear	MethodOfTravel
1	Nicholls	William	Male	7	Walk
2	Brooks	Georgia	Female	9	Bus
3	Shazid	Asif	Male	11	Car
4	Pang	Kwan Li	Female	10	Walk
5	Mulgrew	Alison	Female	9	Walk
6	Isaacs	Martin	Male	9	Cycle
7	Stewart	Gary	Undisclosed	8	Bus
8	Pine	Caroline	Female	11	Bus
9	Shatner	Franklin	Male	10	Car
10	Naeem	Shayla	Female	7	Bus

Explain how you would use this database to obtain a list of all female pupils in Year 8 from this database, with the pupils' names in alphabetical order.

(4)

(Total 4 marks)

---

4. The following spreadsheet shows the number of people working in the public sector between 2000 and 2009, categorised by public sector job type.

Number of public sector workers in the United Kingdom (thousands)

	A	B	C	D	E	F	G	H	I
1						United Kingdon (thousands), seasonally adjusted			
2		HM Forces	Police (including civilians)	Public administration	Education	National Health Service	Other Health and Social Work	Other public sector	Total public sector
3	**2000**	217	218	994	982	990	274	615	4290
4	**2001**	214	222	998	997	1025	265	642	4363
5	**2002**	214	230	1020	1007	1075	263	662	4471
6	**2003**	223	240	1056	1049	1126	249	665	4608
7	**2004**	218	254	1069	1076	1179	260	672	4728
8	**2005**	210	262	1081	1090	1221	267	686	4817
9	**2006**	204	264	1077	1121	1224	265	657	
10	**2007**	197	272	1062	1132	1220	264	664	4811
11	**2008**	193	275	1025	1143	1245	262	679	4822
12	**2009**	197	283	1015	1164	1304	297	960	5220

[Source: www.ons.gov.uk/employmentandlabourmarket/peopleinwork/publicsectorpersonnel/]

(a) Explain how spreadsheet software can be used to calculate the value in cell I9.

(2)

(b) A similar spreadsheet to the one shown above shows the numbers of public sector workers in Germany between 2000 and 2009.

Explain how you would use spreadsheet software to compare the data for "Education" between the two countries.

(3)

(Total 5 marks)

---

5. The Home Office of the UK government collect information about international migration and store it on a database.

The table below shows part of the database schema.

Field	Data Type	Description
geo_name	Text	Name of UK nation state migrated to
geo_code	Text	Code of UK nation state migrated to
mid_year_to_mid_year	Text	The year interval the information were collected for
gender	Text	'Female', 'Male' or 'All'
age	Number	Age of the migrants
net_migration	Number	The difference between the number of migrants entering and the number of migrants leaving.

(a) Explain how you would use this database to obtain a list detailing the migration of females aged 18 and under in Northern Ireland in the year 2018/2019.

(5)

(b) Explain how you would obtain a list of UK Nation codes with no duplication.

(2)

(Total 7 marks)

---

6. The following is a list of Ali's lap times for a selection of 8 tracks in *Mario Kart 8 Deluxe*:

	A	B	C	D	E
1		**Lap 1**	**Lap 2**	**Lap 3**	**Total**
2	**Sunshine Airport**	48.352	46.424	45.19	**139.966**
3	**Dolphin Shoals**	45.502	43.62	43.976	**133.098**
4	**Electrodrome**	48.342	45.049	46.123	**139.514**
5	**Mount Wario**	30.533	51.767	39.066	**121.366**
6	**Cloudtop Cruise**	47.234	45.714	45.564	**138.512**
7	**Bone Dry Dunes**	45.563	44.138	44.145	**133.846**
8	**Bowser's Castle**	47.393	46.848	45.919	**140.16**
9	**Rainbow Road**	48.526	46.402	44.668	**139.596**
10					

(a) Explain how to use spreadsheet software to calculate the expected total time per track?

(2)

(b) Explain how to use spreadsheet software to calculate the spread of all lap times?

(2)

(c) The "Total" column was calculated using a formula.

What formula was used for the total lap time for *Rainbow Road*?

(1)

(Total 5 marks)

---

7. A warehouse stores products that are waiting to be delivered to supermarkets. The products can be identified by labelling them individually with barcodes.

Each barcode stores a representation of a Product ID number that uniquely identifies the type of product, with the details of the product stores on the warehouse database.

The table below shows an excerpt from this database:

Product_ID	Description	Ret_Category	Ret_Subcategory	Quant_In_Stock
102546	Washing Powder 1 kg box	Household	Laundry	10000
398352	Baked Beans 455 g tin	Food	Tinned Produce	1450
293820	Large dishcloths (5 pack)	Household	Kitchen	300

(a) Explain how you would use the database to obtain a list of the descriptions of all items in the "Food" category in alphabetical order.

(3)

(b) Explain how you would use the database to find the total number of items in stock in the "Household" category, "Laundry" subcategory.

(4)

(Total 7 marks)

8. The spreadsheet printout below shows the breakdown of waste from different areas.

	A	B	C	D	E	F
1	**District**	**Population**	**Household waste per person (kg)**	**Household - total waste (tonnes)**	**Non-household - total waste (tonnes)**	**Total local authority waste (tonnes)**
2	**Area 1**					
3	Malvern Hills	81,046	313	25,327	2,857	28,184
4	Worcestershire	610,290	427	260,533	36,542	297,075
5	Worcester City	104,208	302	31,450	2,703	34,153
6	**TOTALS**	**795,544**		**317,310**	**42,102**	**359,412**
7	**Area 2**					
8	Bromsgrove	100,374	364	36,486	5,749	42,235
9	Redditch	87,145	311	27,102	213	27,316
10	Wyre Forest	102,321	342	34,953	2,759	37,712
11	Wychavon	134,832	358	48,270	4,573	52,843
12	**TOTALS**	**424,672**		**146,811**	**13,294**	**160,106**
13	**Area 3**					
14	Herefordshire	188,943	363	68,662	10,200	
15	Shropshire	327,893	479	157,159	8,624	165,782
16	**TOTALS**	**516,836**		**225,821**	**18,824**	**244,643**

[Source: https://www.gov.uk/government/statistics/local-authority-collected-waste-management-annual-results]

The spreadsheet was extracted from data in a database with the following structure:

Name	Type	Width	Decimal	Label
Area_Number	Numeric	8	0	1 (Worcestershire), 2 (Worcestershire), 3 (Borders)
Distict	String	16	0	Name of District
Population	Numeric	8	0	Population size of district
Household_total	Numeric	8	0	Total household waste (tonnes)
Non_household_total	Numeric	8	0	Total non-household waste (tonnes)

(a) (i) Explain how you can use the spreadsheet to calculate the total tonnage of waste from Herefordshire.

(1)

(ii) Explain how you can use the spreadsheet to calculate how many people altogether lived in all three areas?

(1)

(iii) Explain how you can use the spreadsheet to calculate the percentage of these people lived in Area 3?

(2)

(b) Explain how you would use the database to obtain a list of districts which have a total waste of more than 40,000 tonnes.

(3)

(Total 7 marks)

9. The spreadsheet below shows the ethnicity by age group and gender in the UK in 2023 from the Annual Population Survey carried out by the Office of National Statistics.

	A	B	C	D	E	F
1	**Ethnicity by age group and sex, UK, 2023**					
2						
3						*Millions*
4			**16-19**	**20-24**	**25-49**	**50+**
5	**Male**	White	1142.8	1621	8707.7	11370.8
6		Mixed	58.8	73.1	175.3	59.4
7		Indian	36	70.3	492.5	207.9
8		Pakistani/Bangladeshi	72.6	79.9	375	123.6
9		Black	102.7	85.9	402.5	241
10		Other Ethnic Group	67.2	86.8	607	235.9
11	**Female**	White	1119.4	1534	8540.5	12430.5
12		Mixed	49.8	62.9	215.5	74.6
13		Indian	34.5	54.7	454	214.8
14		Pakistani/Bangladeshi	58	77.7	367.8	105.6
15		Black	78.9	91.5	521.1	314.7
16		Other Ethnic Group	70.5	120.4	694.8	272.8
17						
18						
19						*Percentages*
20			**16-19**	**20-24**	**25-49**	**50+**
21	**Male**	White	77.2	80.4	80.9	92.9
22		Mixed	4	3.6	1.6	0.5
23		Indian	2.4	3.5	4.6	1.7
24		Pakistani/Bangladeshi	4.9	4	3.5	1
25		Black	6.9	4.3	3.7	2
26		Other Ethnic Group	4.5	4.3	5.6	1.9
27	**Female**	White	79.3	79	79.1	92.7
28		Mixed	3.5	3.2	2	0.6
29		Indian	2.4	2.8	4.2	1.6
30		Pakistani/Bangladeshi	4.1	4	3.4	0.8
31		Black	5.6	4.7	4.8	2.3
32		Other Ethnic Group	5	6.2	6.4	2

[Source: www.nomisweb.co.uk/home/release_group.asp?g=16]

This information is extracted from a database table called “APS” with the following schema:

ID_Number	Number	Unique number to identify each person
Gender	Text	Gender of person by self-identification
Age	Number	Age of person
Ethnicity	Text	Ethnicity of person

(a) Explain how the database could be used to obtain the value in cell E7

(6)

(b) Explain how the spreadsheet software can be used to obtain the value in cell E21

(2)

The data in cells E21 to E26 are calculated correctly but do not sum to 100.

(c) Give a reason for this discrepancy.

(1)

(Total 9 marks)

TOTAL FOR PAPER: 50 MARKS