EXAMINATIONS OF THE HONG KONG STATISTICAL SOCIETY

HIGHER CERTIFICATE IN STATISTICS, 2016

MODULE 1 : Data collection and interpretation

Time allowed: One and a half hours

Candidates should answer **THREE** *questions.*

Each question carries 20 marks.
The number of marks allotted for each part-question is shown in brackets.

Graph paper and Official tables are provided.

Candidates may use calculators in accordance with the regulations published in the Society's "Guide to Examinations" (document Ex1).

The notation log *denotes logarithm to base* ***e****.*
Logarithms to any other base are explicitly identified, e.g. $\log_{10}$.

Note also that $\binom{n}{r}$ *is the same as* ${}^{n}C_{r}$.

This examination paper consists of 8 printed pages.
This front cover is page 1.
Question 1 starts on page 2.

There are 4 questions altogether in the paper.

1. The table below shows two measures of inequality of income for various countries. For each measure, larger numbers indicate greater inequality.

- The Gini index takes values from 0 to 1. If everyone in a country had exactly the same income, the Gini index would be 0. If all the income went to just a few individuals, the Gini index would be close to 1.
- The 10 / 10 ratio is calculated as the average disposable income of the richest 10% of the population divided by the average disposable income of the poorest 10% of the population.

	Gini Index						**10 / 10 ratio**		
Year	*1985*	*1990*	*1995*	*2000*	*2005*	*2010*	*1995*	*2005*	*2010*
Anglo-Saxon countries									
United States	0.34	0.35	0.36	0.36	0.38	0.38	12.5	15.5	15.1
United Kingdom	0.32	0.37	0.35	0.36	0.33	0.34	7.1	8.9	10.1
Australia			0.30	0.31	0.31	0.34	7.7	7.7	9.3
Canada	0.29	0.29	0.29	0.32	0.32	0.32	7.2	8.8	9.0
Northern Europe									
Finland		0.22	0.22	0.25	0.27	0.26	5.4	6.6	5.9
Sweden	0.21	0.21	0.21	0.24	0.24	0.26	4.1	4.7	5.8
Denmark	0.22	0.23	0.21	0.23	0.23	0.25	4.2	5.2	5.5
Norway	0.22	0.23	0.24	0.25	0.28	0.25	4.0	4.6	5.2
Western Europe									
Portugal		0.33	0.36	0.36	0.37	0.35	10.5	12.4	10.1
Italy	0.31		0.35	0.34	0.35	0.34	11.5	10.7	9.7
Germany	0.25	0.26	0.27	0.26	0.30	0.30	5.3	6.6	7.2
The Netherlands	0.26	0.29	0.28	0.29	0.28	0.29	6.0	6.6	7.1
Austria	0.24		0.24	0.25	0.26	0.26	4.9	6.0	5.6
Other countries									
Mexico	0.45	0.50	0.50	0.51	0.49	0.48	33.5	25.8	27.0
Japan	0.30		0.32	0.34	0.33		10.2	10.1	10.3

Write a report summarising the information in the table for an intelligent but non-specialist audience. Your report should look at

- differences within and between groups of countries,
- trends over time,
- whether or not the Gini index and the 10 / 10 ratio give a consistent message.

You should incorporate diagrams into your report where appropriate.

(20)

2. In the United Kingdom there has been a national census every 10 years since 1801 (with the exception of 1941). At the time of the 2011 UK census, a government minister described the census as 'expensive, inaccurate and inefficient', and 'out of date almost before it's done'.

The minister also said that data held by the National Health Service, local councils, the postal service, the electoral register, tax returns, credit card firms and phone companies can do the job.

A proposal for the 2021 UK census is that it should be conducted online and that it should incorporate additional data held by government agencies.

(i) Describe the purpose of a census, explaining how it differs from a sample survey and from routine collection of data by government agencies.

(5)

(ii) The 2011 UK census attracted a response rate of about 94% of the population. Discuss whether or not that is a problem for the accuracy of the census.

(4)

(iii) In the 2011 UK census, almost 170 000 people stated their religion as 'Jedi Knight'. (Jedi Knights are characters in the 'Star Wars' films.)

Discuss what responses of this type indicate about the attitudes of some members of the public to the census. Discuss also whether responses of this type invalidate asking a question about religion.

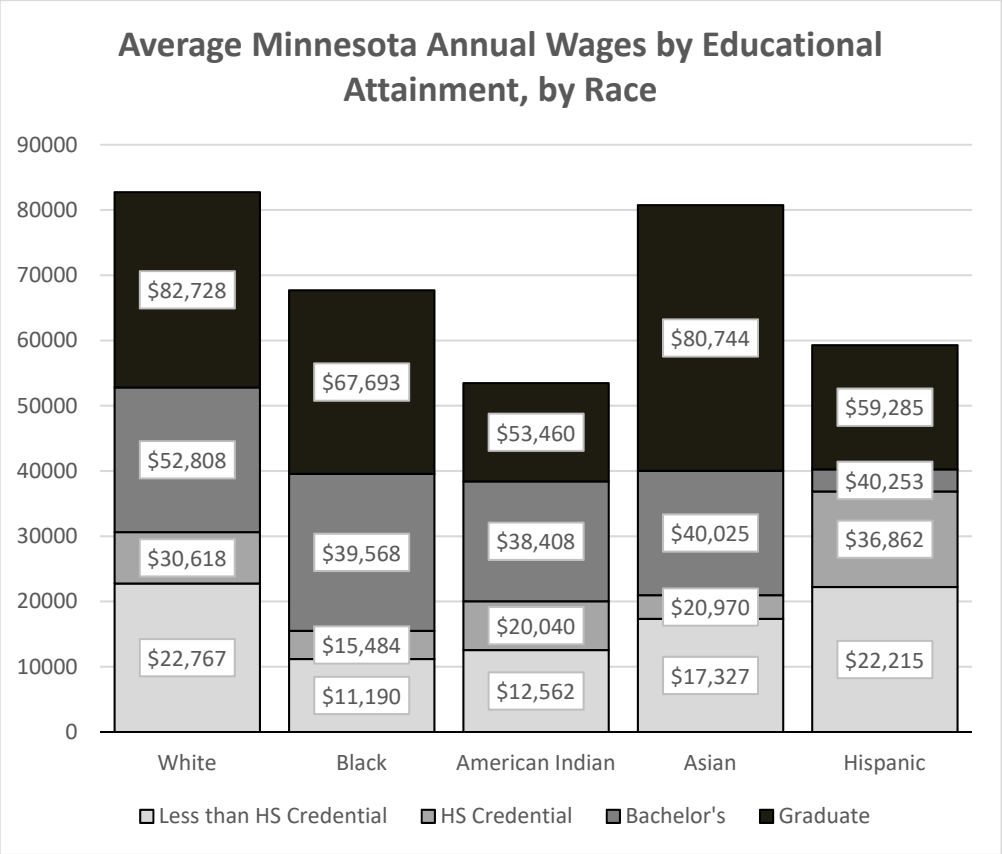
(6)

(iv) Discuss the potential problems in conducting the 2021 UK census online, and explain how these problems might be overcome.

Discuss the potential problems in incorporating additional data held by government agencies.

(5)

3. (a) The chart below shows, for a recent year, average annual wages in Minnesota for people of different levels of educational achievement and different race. ('HS credential' indicates that someone has completed high school.)



Identify the strengths and weaknesses in the way that the data have been presented. State what additional data would be useful, and explain how you would present the information if those were available.

(10)

(b) The chart **on the next page** shows, for the OECD countries, the obesity rates of adults with different levels of educational achievement.

(i) Identify the strengths and weaknesses in the way the data have been presented. Where appropriate, suggest ways in which the presentation could be improved.

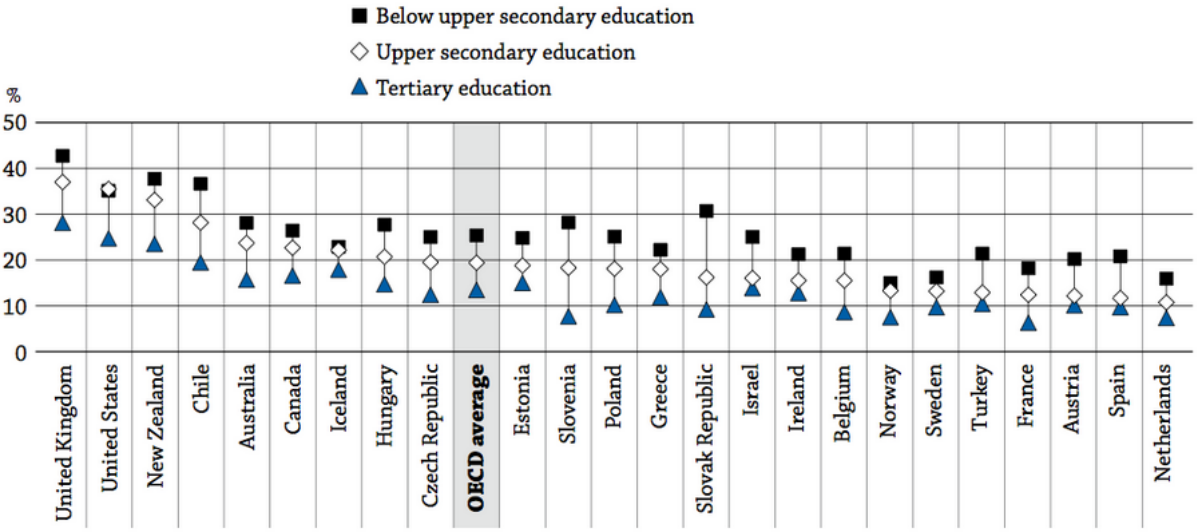
(5)

(ii) Summarise in words the main messages in these data.

(5)

Chart for Question 3(b) is on the next page

Proportion of obese adults, by level of educational attainment (2011)



4. The data in the table are taken from a survey of the diet of 2848 adults in a region of China.

The numbers of men and women surveyed were divided, separately, into four equal parts on the basis of their fresh vegetable consumption. These parts are shown as Q1, Q2, Q3 and Q4. The mean consumption in grams per day is given for each food type in each part, for men and women separately. For example, the mean consumption of fruit by men who ate the lowest amount of fresh vegetables is 31 grams per day.

The columns headed SE give the standard errors of the mean food consumptions by all men and all women.

	Men (sample size 1308)					*Women (sample size 1540)*				
	Q1	Q2	Q3	Q4		Q1	Q2	Q3	Q4	
	Mean	Mean	Mean	Mean	SE	Mean	Mean	Mean	Mean	SE
Fresh veg	204	259	266	317	0.9	178	235	266	304	0.8
Fruit	31	45	69	105	0.5	28	46	70	121	0.4
Rice	367	337	269	246	1.0	315	276	243	220	0.8
Wheat flour	79	114	197	253	1.0	56	118	141	180	0.8
Whole grain	2	2	6	27	0.1	1	3	6	22	0.1
Root veg	7	11	16	29	0.1	6	12	17	28	0.1
Meat	70	61	69	77	0.4	48	43	54	63	0.3
Fish	23	28	31	44	0.2	19	21	28	46	0.2
Milk	2	3	23	39	0.3	1	4	15	48	0.3

(i) Calculate the overall mean consumption of fresh vegetables for men and women separately. Give the underlying standard deviation in each case.

Calculate also the overall mean consumption of fresh vegetables for men and women combined.

(6)

(ii) Describe in words what the figures for milk consumption indicate.

(3)

(iii) What distinctive pattern is there, for both men and women, in rice consumption across the four parts, Q1 to Q4?

(2)

(iv) Draw a suitable diagram to show the difference in consumption patterns of fish between men and women.

(4)

(v) Men require, on average, about 20% more food per day than women to maintain energy levels. Use this information to compare the consumptions of the main food groups by men in Q1 and women in Q1. (You are not required to consider whole grains, root vegetables or milk.)

(5)

BLANK PAGE

BLANK PAGE