

Stockholms Universitet, Statistisk

Exam in: Statistical databases and register

Examiner: Mikael Möller and Bo Sundgren

Approved aids: Pen and pencil

Examination day: 100312

Examination room: B705

Examination time: 3 hours

All assumptions and notations should be explained and defined (also those that have been used during the course). All answers, reasoning and explanations should be easy to follow. Answers and arguments which cannot be understood give 0 p.

Good luck!

1: You are interested in statistics about commuters. Each commuter live and work in different communes. The following information is available: (10p)

- **about commuter:** id, age, sex, education, occupation, home commune, link to establishment of work (an establishment is a part of a company/organisation located at a certain place)
- **about company/organisation:** id, links to the establishments of the company/organisation (a company/organisation may have one or more establishments)
- **about establishment:** location (commune), number of employees, link to the company/organisation to which the establishment belongs

- a. Draw an object graph (conceptual model) for this area of interest, showing object types, variables of the object types, and relations between object types.
 - b. Transform the object graph into a relational data model, consisting of relational tables with rows and columns, primary keys (identifiers) and secondary keys (links to other relations tables in the data model).
- 2:** Referring to the previous exercise. One statistic of interest is the following one:
- **number of commuters** between home commune of commuter and location of establishment where the commuter works, age of commuter and sex of commuter.

- a. This statistic could be described by means of a four-dimensional hypercube. Specify the dimensions and the contents of the cells of this hypercube. What is the population of objects associated with the hypercube? (5p)
- b. Alternatively the statistic could be described by means of a so-called star model. How? (5p)

This problem may also be formulated to the case of the **number of travels** between the communes. You get extra points (max 5p) if you also regard this case.

3: During the course it was shown how the types of objects occurring in official statistics roughly could be categorised into three major types: actors, utilities, and complex objects connecting (relationships, activities, events, etc) actors and utilities. It was also shown how models for official statistics could be drawn in a standardised way: actors appearing to the left, utilities to the right and the complex objects in the middle. (10p)

- a. Take a second look at the model you draw in task 1a. Has it been drawn in the standard just mentioned? If not, draw another version of the graph, which follows the standard: actors to the left, utilities to the right, and complex objects in the middle.
- b. Which objects in your graph are actors, utilities and complex objects? Give a formal definition of the object actor/utility/complex (all three)!

4: Study the table *Befolkningsutveckling 1749-2009* (an excerpt from SCB) on the next page.

- a. Which parts of what you see on the page, would you define as “data”, and what would you define as “metadata”? (3p)
- b. List a number of important users/usages of statistical metadata! (3p)
- c. Which kinds of metadata are particularly important for the respective metadata users/usages that you have just listed? (2p)
- d. You have just been appointed chief designer and project leader for a metadata project in your organisation. List a number of things that you should think about in order to succeed with the project! (2p)

År	Folkmängd	Födda	Döda	Invandrare	Utvandrare	Giftermål	Skilsmässor
1749	1764724	59483	49516	15046
1750	1780678	64511	47622	16374
1751	1802132	69291	46902	16599
.....							
1962	7581148	107284	76791	25084	14928	53913	8849
1963	7627507	112903	76460	26950	15340	53480	8496
1964	7695200	122664	76661	38334	15705	58439	9169
1965	7772506	122806	78194	49586	15977	59963	9563
1966	7843088	123354	78440	46970	19730	61101	10288
1967	7892774	121360	79783	29983	19979	56561	10722
1968	7931193	113087	82476	35978	23162	52291	11228
1969	8004270	107622	83352	64503	20360	48357	12140
1970	8081229	110150	80026	77326	28653	43278	12943
1971	8115165	114484	82717	42615	39560	39918	13540
1972	8129129	112273	84056	29894	41579	38636	15179
1973	8144428	109663	85640	29443	40342	38251	16021
1974	8176691	109874	86316	37430	28352	44864	26802
1975	8208442	103632	88208	44133	27249	44103	25383
1976	8236179	98345	90677	45492	25522	44790	21702
1977	8267116	96057	88202	44005	21078	40370	20391
1978	8284437	93248	89681	36187	22168	37844	20317
1979	8303010	96255	91074	37025	23467	37300	20322
1980	8317937	97064	91800	39426	29839	37569	19887
1981	8323033	94065	92034	32272	29440	37793	20198
1982	8327484	92748	90671	30381	28381	37051	20766
1983	8330573	91780	90791	27495	25269	36210	20618
1984	8342621	93889	90483	31486	22825	36849	20377
1985	8358139	98463	94032	33127	22036	38297	19763
1986	8381515	101950	93295	39487	24495	38906	19107
1987	8414083	104699	93307	42666	20673	41223	18426
1988	8458888	112080	96743	51092	21461	44229	17746
1989	8527036	116023	92110	65866	21484	108919	18862
1990	8590630	123938	95161	60048	25196	40477	19357
1991	8644119	123737	95202	49731	24745	36836	20149
1992	8692013	122848	94710	45348	25726	37173	21907
1993	8745109	117998	97008	61872	29874	34005	21673
1994	8816381	112257	91844	83598	32661	34203	22237
1995	8837496	103422	93955	45887	33984	33642	22528
1996	8844499	95297	94133	39895	33884	33784	21377
1997	8847625	90502	93326	44818	38543	32313	21009
1998	8854322	89028	93271	49391	38518	31598	20761
1999	8861426	88173	94726	49839	35705	35628	21000
2000	8882792	90441	93461	58659	34091	39895	21502
2001	8909128	91466	93752	60795	32141	35778	21022
2002	8940788	95815	95009	64087	33009	38012	21322
2003	8975670	99157	92961	63795	35023	39041	21130
2004	9011392	100928	90532	62028	36586	43088	20106
2005	9047752	101346	91710	65229	38118	44381	20000
2006	9113257	105913	91177	95750	44908	45551	20295
2007	9182927	107421	91729	99485	45418	47898	20669
2008	9256347	109301	91449	101171	45294	50332	21377
2009	9340682	111803	90080	102280	39240	48033	22211