Major in Statistics
Bachelor of Science
(4-year Curriculum)

What is STATISTICS about?

Data / Observations

Statistical methods

Inference about "TRUTH" / Decision-making

Applications

whenever you want to study data…

i.e.   *in nearly all disciplines!*

Gold Medalist
Louganis, Gregory

GAMES OF THE XXIVTH OLYMPIAD SEOUL 1988

Beijing 2008
TM©

Gold Medalist
Mitcham, Matthew

Silver Medalist
Xiong, Ni (熊倪)

Silver Medalist
Zhou, Luxin (周吕鑫)

Results of final (10th round) dive

Men Platform Diving
Final Dive 10/Sept. 27

Hommes Plongeon de Haut-vol
Finale Plongeon 10/27 Sept

Order Ordre	Rank Rang	Ctry Pays	Name	Type Saut	Diff Coef	Judge / Juge							Total
						1	2	3	4	5	6	7	
1	9	JPN	Yamagishi, Isao	405 B	2.8	6.0	6.0	7.0	7.0	6.5	7.0	6.5	55.44
2	12	MEX	Mondragon, Jorge	307 C	3.4	3.0	4.0	3.5	4.0	4.0	4.5	4.5	40.80
3	3	MEX	Mena, Jesus	5237D	3.2	8.0	7.5	7.5	7.5	8.0	7.5	8.0	73.92
4	7	CAN	Bedard, David	5237D	3.2	7.0	6.0	6.0	7.0	6.5	6.0	6.5	61.44
5	8	GDR	Haage, Steffen	5237D	3.2	6.0	6.5	6.0	5.5	6.5	6.0	6.0	58.56
6	4	URS	Tchogovadze, Gueorgui	407 C	3.2	7.0	7.5	7.5	7.5	8.0	7.5	7.0	71.04
7	10	USA	Jeffrey, Patrick	407 C	3.2	5.5	5.5	4.5	6.0	6.0	5.5	5.5	53.76
8	6	GDR	Hempel, Jan	5237D	3.2	6.5	7.5	6.0	6.5	7.0	6.5	6.5	63.36
9	5	URS	Timochinine, Vladimir	5237D	3.2	7.0	7.5	7.0	7.0	7.0	7.0	7.0	67.20
10	11	CHN	Li, Kongzheng	5337D	3.3	5.0	6.0	5.5	5.0	4.5	6.5	5.5	53.46
11	2	CHN	Xiong, Ni	407 C	3.2	8.5	8.5	9.0	9.0	7.0	8.5	8.5	82.56
12	1	USA	Louganis, Gregory	307 C	3.4	8.5	8.5	8.5	8.5	8.0	8.5	9.0	86.70

12 finalists

7 judges

Results of final (6th round) dive

highest in Olympic history

NOC	Name	Dive No.	DD	Judge's Score							Dive Points	Dive Rank	Total Points	Overall Rank
				J1	J2	J3	J4	J5	J6	J7				
Australia	MITCHAM Matthew	5255B	3.8	10	9.5	10	10	9	9.5	10	112.1	1	537.95	1
China	ZHOU Luxin	307C	3.4	6.5	7.5	7	7.5	6	8	8	74.8	8	533.15	2
Russian Fed.	GALPERIN Gleb	5255B	3.8	10	8.5	9	9.5	8.5	9	9	102.6	2	525.8	3
China	HUO Liang	5255B	3.8	6	7.5	8.5	7	8	8	7	85.5	5	508.4	4
Cuba	GUERRA OLIVA Jose Antonio	5255B	3.8	8.5	8.5	9	8.5	8.5	9	8.5	96.9	3	507.15	5
Australia	HELM Mathew	5253B	3.4	8.5	8.5	8.5	8.5	8.5	9	9	86.7	4	467.7	6
Great Britain	DALEY Thomas	307C	3.4	5.5	6.5	6	6	7.5	6.5	7.5	64.6	10	463.55	7
Mexico	PACHECO Rommel	5253B	3.4	8	8	8	8	8.5	8	8.5	81.6	7	460.2	8
Germany	HAUSDING Patrick	5255B	3.8	6	6	6.5	6.5	6.5	7	6.5	74.1	9	448.3	9
United States	BOUDIA David	5255B	3.8	3.5	3.5	4	4	4	4.5	3.5	43.7	12	441.45	10
Colombia	URAN Juan Guillermo	207B	3.6	4	4.5	5	4.5	5	4	4.5	48.6	11	414.8	11
United States	FINCHUM Thomas	5253B	3.4	8.5	8	8.5	8.5	8	8	8.5	85	6	412.65	12

12 finalists

7 judges

How fair were the Olympic judges?

7 judges

Any judge biased towards/against any finalist?

12 finalists

Major in Decision Analytics
Bachelor of Science
(4-year Curriculum)

What is DECISION ANALYTICS about?

Big Data

Data Mining / Statistical Learning / Machine Learning

Decision-making

Applications

- **Marketing**
  - *e.g. customer preference / behaviour / loyalty*
- **Finance**
  - *e.g. customer banking, stock volatility analysis, relationships between financial indicators*
- **Forensic Accounting**
  - *e.g. fraud detection in credit card transactions / insurance claims, tax evasion, insider-trading operations*
- **Healthcare**
  - *e.g. risk factors, treatment effectiveness, medical planning*
- **Biology**
  - *e.g. biological function via gene expression, protein structure identification / prediction*
- plus many others…

Netflix movie-rating challenge

Movie-rental company

17,770 movies

480,189 customers

	Dirty Dancing	Meet the Parents	Top Gun	The Sixth Sense	Catch Me If You Can	The Royal Tenenbaums	Con Air	Big Fish	The Matrix	A Few Good Men
Customer 1					4					
Customer 2			3				3			3
Customer 3		2		4					2	
Customer 4	3									
Customer 5	5	5			4					
Customer 6						2	4			
Customer 7			5					3		
Customer 8						2				3

How to predict missing ratings of customers?

Netflix Prize

COMPLETED

Home | Rules | Leaderboard | Update

Leaderboard

Showing Test Score. Click here to show quiz score
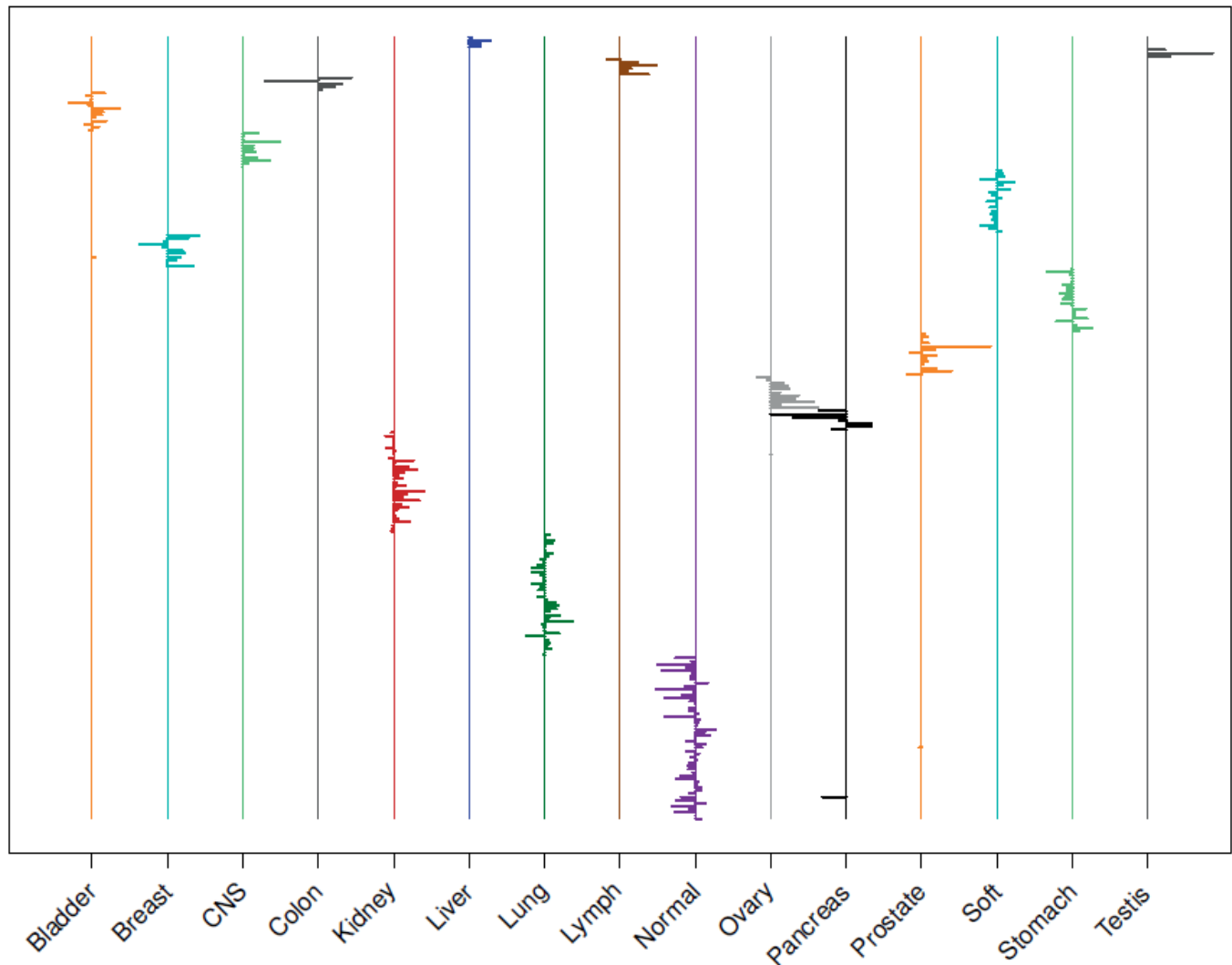
Display top 20 ✓ **leaders.**

… involved many statistical techniques, of which most important is SVD (Singular Value Decomposition)

Rank	Team Name	Best Test Score	% Improvement	Best Submit Time
Grand Prize - RMSE = 0.8567 - Winning Team: BellKor's Pragmatic Chaos				
1	BellKor's Pragmatic Chaos	0.8567	10.06	2009-07-26 18:18:28
2	The Ensemble	0.8567	10.06	2009-07-26 18:38:22
3	Grand Prize Team	0.8582	9.90	2009-07-10 21:24:40
4	Opera Solutions and Vandelay United	0.8588	9.84	2009-07-10 01:12:31
5	Vandelay Industries !	0.8591	9.81	2009-07-10 00:32:20
6	PragmaticTheory	0.8594	9.77	2009-06-24 12:06:56
7	BellKor in BigChaos	0.8601	9.70	2009-05-13 08:14:09
8	Dace_	0.8612	9.59	2009-07-24 17:18:43
9	Feeds2	0.8622	9.48	2009-07-12 13:11:51
10	BigChaos	0.8623	9.47	2009-04-07 12:33:59
11	Opera Solutions	0.8623	9.47	2009-07-24 00:34:07
12	BellKor	0.8624	9.46	2009-07-26 17:19:11

Gene expression cancer data

Cancer type ⬅ **4,718 genes**

Cancer type ⬅ **4,718 genes**

⋮

349 patients

What genes characterise which type of cancer?
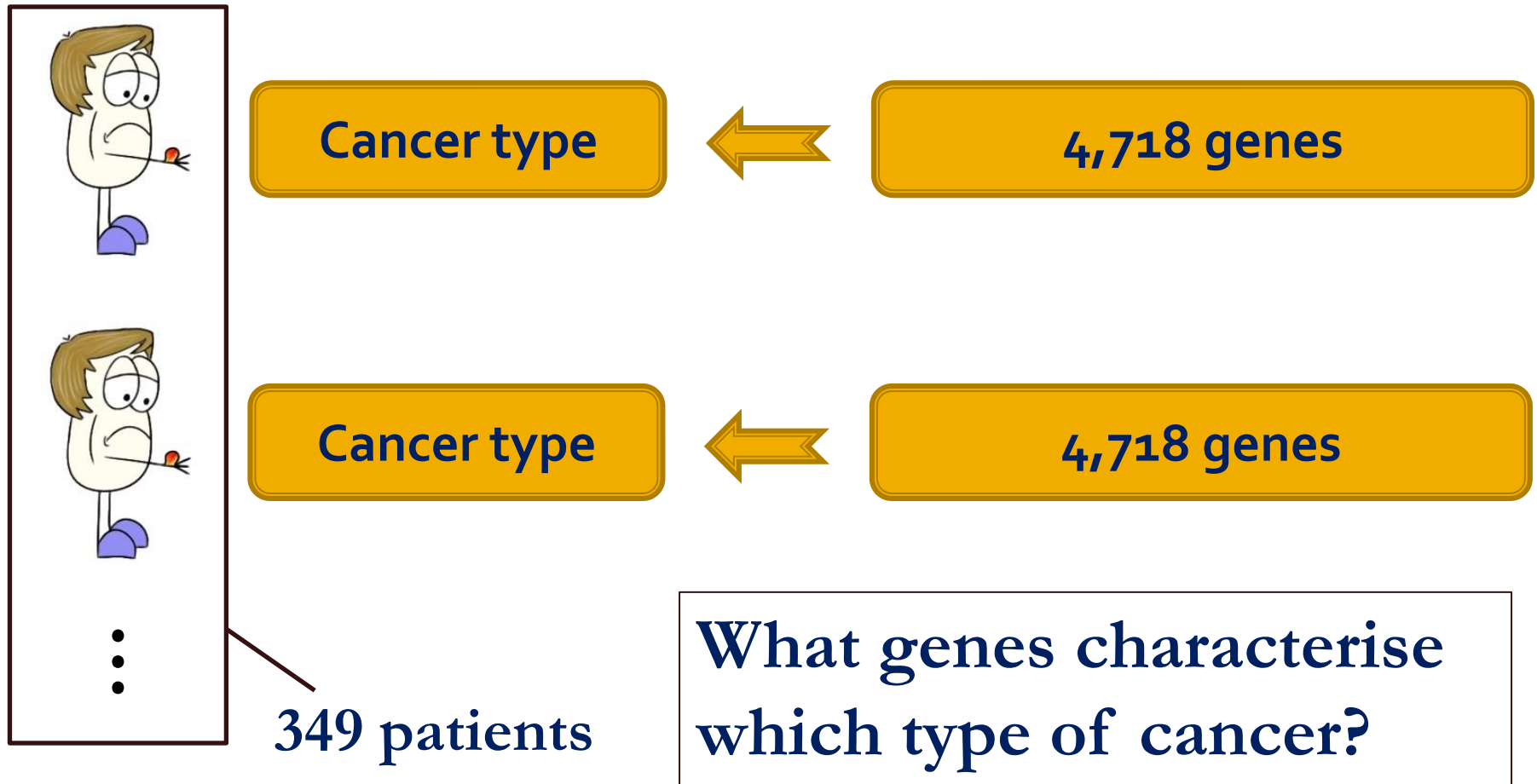
Using lasso-regularised multinomial classifier…

genes

Bladder Breast CNS Colon Kidney Liver Lung Lymph Normal Ovary Pancreas Prostate Soft Stomach Testis

Major in Statistics
Bachelor of Science
(4-year Curriculum)

Structure of curriculum…

7 introductory level courses (42 credits)

- SCNC1111 Scientific method & reasoning

- SCNC1112 Fundamentals of modern science

- *MATH1013 University mathematics II

- *MATH2014 Multivariable calculus & linear algebra

- ***STAT1600**   **Statistics: ideas & concepts**

- ***STAT2601**   **Probability and statistics I**

- ***STAT2602**   **Probability and statistics II**

replaced by other advanced level STAT course(s) if already taken to fulfill other majors/minors

Mathematical background adequate?

- Students must have <u>level 2 or above</u> in HKDSE Extended Module 1 or 2 of Mathematics
or equivalent

- Otherwise, strongly advised to take **MATH1011 University Mathematics I** in Semester 1.

What do we need from your Mathematics?

- Set notation and theory
- Functions (*incl.* limits, continuity)
- Sequences, series
- Basic calculus (*incl.* partial differentiation, double integration)
- Vectors, matrices (basic operations)

Suggested / Example Structure of BSc (Major in Statistics) Curriculum

Year	One		Two	
Semester	One	Two	One	Two
Disciplinary Core	MATH1013 University Mathematics II STAT1600 Statistics: Ideas and Concepts	MATH2014 Multivariable Calculus and Linear Algebra	STAT2601 Probability and Statistics I	STAT2602 Probability and Statistics II
Science Foundation Courses	SCNC1111 Scientific Method and Reasoning	SCNC1112 Fundamentals of Modern Science		
Common Core	Six common core courses within the first three years			
Language	CAES1000 Core University English (offered in both semesters)		CAES9820 Academic English for Science Students (offered in both semesters)	

4 Core advanced level courses (24 credits)

> **Also core to Decision Analytics Major**

STAT3600	Linear statistical analysis
STAT4601	Time-series analysis

> **Also core to Risk Management Major**

STAT3603	Probability modelling
STAT4602	Multivariate data analysis

Other advanced level courses (24 credits = 4 courses)

- **At least one course from *List A*:**

  - **STAT3602 Statistical inference**

  - **STAT3604 Design and analysis of experiments**

  - **STAT3620 Modern nonparametric statistics**

  - **STAT3621 Statistical data analysis**

  mainstream statistics courses

  - *core knowledge in statistics*

  - *applications to general problems in all areas*

  - *foundation for graduate studies in statistics or related subjects*

- **Other courses from *List B*:**
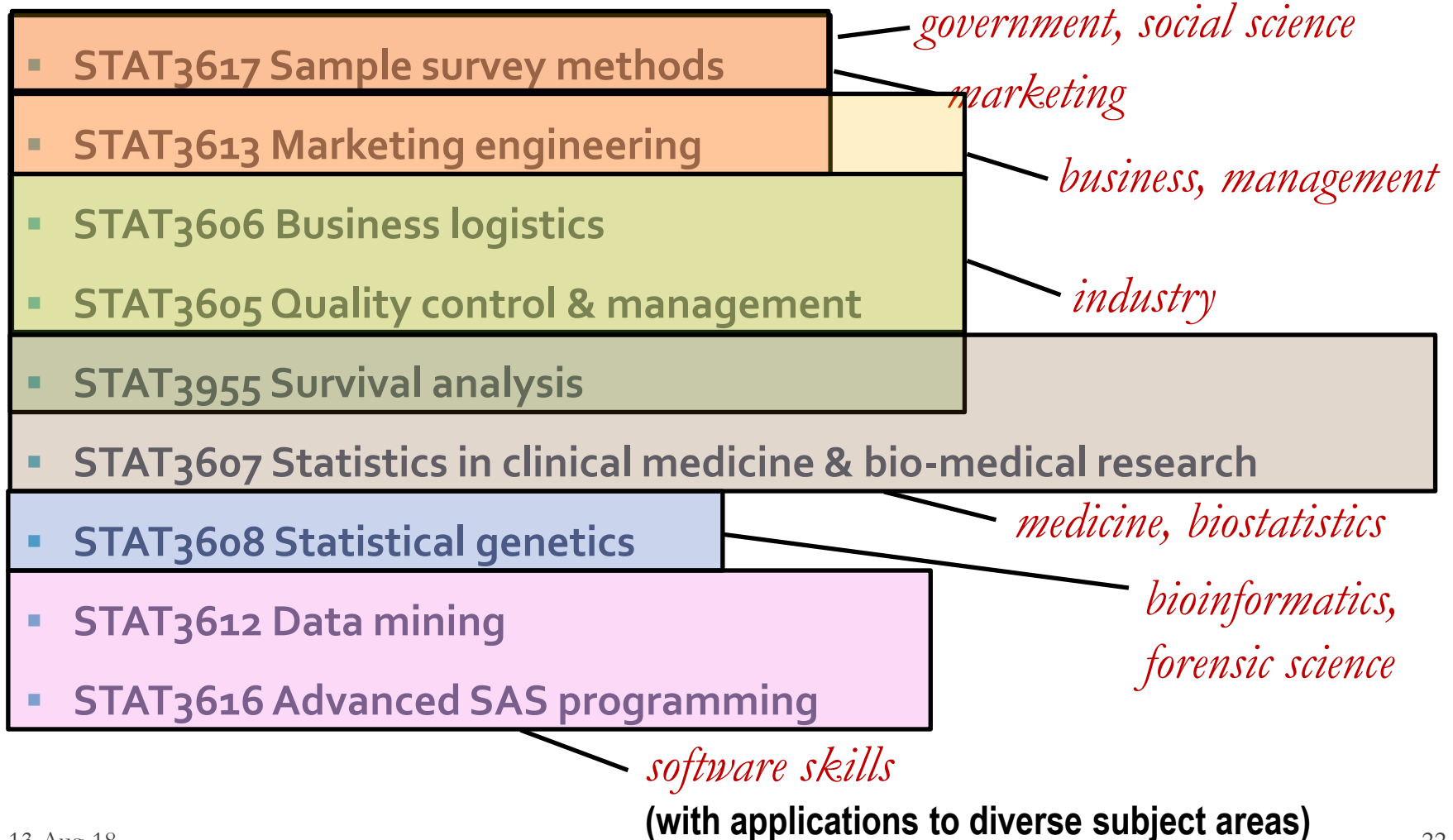
  - STAT3617 Sample survey methods
  - STAT3613 Marketing engineering
  - STAT3606 Business logistics
  - STAT3605 Quality control & management
  - STAT3955 Survival analysis
  - STAT3607 Statistics in clinical medicine & bio-medical research
  - STAT3608 Statistical genetics
  - STAT3612 Data mining
  - STAT3616 Advanced SAS programming

  government, social science

  marketing

  business, management

  industry

  medicine, biostatistics

  bioinformatics, forensic science

  software skills

  (with applications to diverse subject areas)

Other advanced level courses (24 credits = 4 courses)

- **Other courses from *List B*:**

  - STAT3617 Sample survey methods

  - STAT3613 Marketing engineering

  - STAT3606 Business logistics

  - STAT3605 Quality control & management

  conceptually less demanding

  - STAT3955 Survival analysis

  - STAT3607 Statistics in clinical medicine & bio-medical   research

  - STAT3608 Statistical genetics

  - STAT3612 Data mining

  - STAT3616 Advanced SAS programming

  conceptually more demanding

Other advanced level courses (24 credits = 4 courses)

- **Other courses from *List B*:**

  - STAT3617 Sample survey methods
  - STAT3613 Marketing engineering
  - STAT3606 Business logistics
  - STAT3605 Quality control & management
  - STAT3955 Survival analysis
  - STAT3607 Statistics in clinical medicine & bio-medical   research
  - STAT3608 Statistical genetics
  - STAT3612 Data mining
  - STAT3616 Advanced SAS programming

  Theme of data science

  Preparatory course: **STAT2603 Data Management with SAS**

Capstone requirement (6 credits)

At least 6 credits from:

individual project

- STAT3799 Directed studies in statistics (6 credits)

- STAT4799 Statistics project (12 credits)

- STAT4710 Capstone experience for statistics undergraduates (6 credits)

- STAT4766 Statistics internship (6 credits)

group project

Major in Risk Management *vs* Major in Statistics

- All 7 introductory level courses **SAME**
- Advanced level core courses:

  2 SAME,   **2 DIFFERENT**

- **Risk Management** —

  courses focus primarily on business-related topics: e.g.

  investment, insurance, finance, banking, etc.

- **Statistics** —

  courses cover wide range of topics with emphasis on

  "METHODS", their applications, and *underlying theory.*

- Students **CANNOT** double major or major/minor in
  Statistics & **Risk Management**

Major in Decision Analytics
Bachelor of Science
(4-year Curriculum)

Structure of curriculum…

8 introductory level courses (48 credits)

- SCNC1111 Scientific method & reasoning
- SCNC1112 Fundamentals of modern science
- *MATH1013 University mathematics II
- *MATH2014 Multivariable calculus & linear algebra
- *COMP1117 Computer programming
- *COMP2119 Introduction to data structures and algorithms
- ***STAT2601 Probability and statistics I**
- ***STAT2602 Probability and statistics II**

*replaced by other advanced level STAT course(s) if already taken to fulfill other majors/minors

Remarks:

- Students must have <u>level 2 or above</u> in

  HKDSE Extended Module 1 or 2 of Mathematics
  or equivalent;

  otherwise, strongly advised to take
  <u>**MATH1011 University Mathematics I**</u> in Semester 1.
- To enrol in COMP2119, students must also take

  <u>**COMP2123 Programming Technologies and Tools**</u>

Suggested / Example Structure of BSc (Major in Decision Analytics) Curriculum

Year	One		Two	
Semester	One	Two	One	Two
Disciplinary Core	COMP1117 Computer Programming MATH1013 University Mathematics II	MATH2014 Multivariable Calculus and Linear Algebra STAT2601 Probability and Statistics I	STAT2602 Probability and Statistics II	COMP2119 Introduction to Data Structures and Algorithms
Other			COMP2123 Programming Technologies and Tools (Pre-requisite of COMP2119)	
Science Foundation Courses	SCNC1111 Scientific Method and Reasoning	SCNC1112 Fundamentals of Modern Science		
Common Core	Six common core courses within the first three years			
Language	CAES1000 Core University English (offered in both semesters)		CAES9820 Academic English for Science Students (offered in both semesters)	

5 Core advanced level courses (30 credits)

STAT3600   Linear statistical analysis

also core to Statistics / Risk Management Major

STAT3612   Data mining
STAT4609   Big data analytics
COMP3278   Introduction to database management systems
MATH3904   Introduction to optimization

Elective advanced level courses (12 credits = 2 courses)

- **COMP3250 Design & analysis of algorithms**

- **COMP3270 Artificial intelligence**

- **COMP3323 Advanced database systems**

- **COMP3407 Scientific computing**

- **MATH3408 Computational methods & differential equations with applications**

- **MATH3600 Discrete mathematics**

- **MATH3601 Numerical analysis**

- **MATH3901 Operations research**

Elective advanced level courses (12 credits = 2 courses)

- **STAT3616 Advanced SAS programming**
- **STAT3620 Modern nonparametric statistics**
- **STAT3621 Statistical data analysis**
- **STAT3622 Data visualization**
- **STAT4601 Time-series analysis**
- **STAT4602 Multivariate data analysis**

Capstone requirement (6 credits)

At least 6 credits from:

individual project

- STAT3799 Directed studies in statistics (6 credits)

- STAT4799 Statistics project (12 credits)

- STAT4710 Capstone experience for statistics undergraduates (6 credits)

- STAT4766 Statistics internship (6 credits)

group project

Major in Decision Analytics
vs Major in Statistics / Risk Management

- 8 introductory level courses (1 more than STAT/RM)

  - **COMP1117 Computer programming**

  - **COMP2119 Introduction to data structures & algorithms**

    (instead of STAT1600 Statistics: ideas and concepts)

- 5 Advanced level core courses (1 more than STAT/RM):

  one common to all 3 majors: STAT3600 Linear statistical analysis

- Heavier emphasis on

  COMPUTER SCIENCE + MATHEMATICS

- Students **CANNOT** double major in **Decision Analytics** and

  Statistics, Risk Management, Computing & Data Analytics, Computer Science

- Students **CANNOT** major in **Decision Analytics** and minor in

  Statistics, Computer Science

Reminder

- plan ahead

- watch out for pre-requisites of individual courses

- courses **<u>CANNOT</u>** be double-counted to fulfill different majors/minors

(exception for double major in Science:

<u>SCNC1111</u> & <u>SCNC1112</u> & <u>up to 12 credits</u> of compulsory courses

REQUIRED by both Science majors can be double-counted)

- consult course selection advisors if necessary

Contact Persons

- Course Selection Advisors
  - C W Kwan
  - Stephen M S Lee
  - K Zhu
- Tel: 3917 2466
- Email: ug_enquiry@saas.hku.hk

Support from University and Department

- **Department** : Computing facilities
  - two statistical computer laboratories
  - up-to-date software for teaching, learning, research
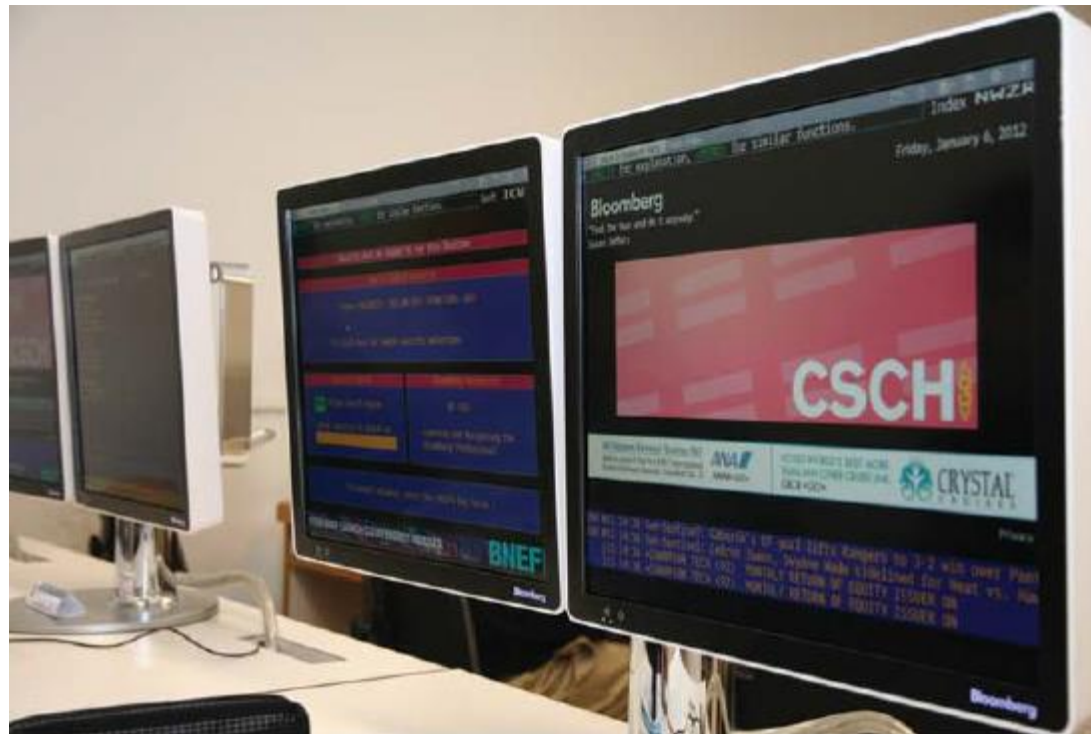
Support from University and Department

- **HKU** : Career Development and Resources for Students
  - NETmatch, NETjobs, JIJIS (Joint Institution Job Info. System)
- **Department** : Internship / Job Online Application System

Contact person:
Dr C W KWAN
<cwkwan@hku.hk>

Career Development Training

- **Summer IT course:**
  - Essential IT skills for statistical and risk analysts
- **Career Advising Programme (CAP) to prepare students for:**
  - internships and job opportunities
  - advancing resume and interview skills

Exchange study

Faculty of Science

Course Equivalence Database

(for credit transfer reference):

http://webapp.science.hku.hk/student/servlet/course_equiv

Contact person:
Dr ZHU Ke
<mazhuke@hku.hk>

Course Equivalence Database

16. Canada - University of British Columbia ▼

Canada - University of British Columbia

No.	Partner Code	Partner Credit/Unit	Partner Title	HKU Code	HKU Credit	HKU Title
1.	BIOC 302	3	General Biochemistry	BIOC2601	6	Metabolism
2.	BIOL 300	3	Fundamentals of Biostatistics	BIOL1608	6	Biostatistics
3.	BIOL 336	3	Fundamentals of Evolutionary Biology	BIOL2210	6	Evolution
					6	Terrestrial Ecology
5.	COMM 370	3	Corporate Finance	STAT2807	6	Corporate Finance for Actuarial Science
6.	COMM 371	3	Investment Theory	STAT3806	6	Investment and Asset Management
7.	COMM 399	3	Logistics and Operations Management	STAT2306	6	Business Logistics
8.	COMM 473	3	Business Finance	STAT2807	6	Corporate Finance for Actuarial Science
9.	COMM 474	3	Fixed Income Markets and Management	STAT3806	6	Investment and Asset Management
10.	MATH 303	3	Introduction to Stochastic	STAT2303	6	Probability Modelling

Scholarships

- Available to students majoring in

  Decision Analytics / Statistics / Risk Management

  based on academic performance and/or other qualities

Student Peer Advisers in 2018-19

- General roles
  - to **offer advice** in relation to academic studies to freshmen; and
  - to **facilitate** freshmen's **smooth transition** from secondary to university education
- You are highly encouraged to contact the following **Student Peer Advisers (SPAs)** if you have any questions about your study (their contacts can be found at the Faculty's website)

Decision Analytics
  - Mr FAN Kwok Lung (BSc Year 4)
  - Mr LAO Annan (BSc Year 3)
  - Miss LUO Tianling (BSc Year 2)
  - Miss ZHENG Shumeng (BSc Year 4)

Statistics
  - Miss LI Ruohan (BSc Year 4)
  - Miss ZHU Zhengyi (BSc Year 2)

Let's talk to our SPAs!

Q & A