Faculty of Information Technology
Department of Information Systems

Approved
Vice-Rector of Academic and Educational Affairs
of IITU JSC
_____ Umarov T.F.
«____» _____ 20__

SYLLABUS
(ACADEMIC PROGRAM)

Course: SP 3304 SAS PROGRAMMING
(code, title)

Major: 5B070300 Information Systems
(code, title)

Educational program 6B06105 «Information System» **(IS, BDA, BA)**
(code, title)

Year: 3, **Semester**: 5; **Number of credits**: 5

Lectures: 15 hours

Laboratory classes: 30 hours

Practical classes: 15 hours

T/SIS: 105 hours

Total: 150 hours

Final assessment form: Project

Almaty 2020

Academic Program of the course «SAS programming»
(course name)

has been developed on the basis of a Standard Academic program.

Academic program has been reviewed at the meeting of «Information system»
Department name

Minutes № 1 dated «17» 08 2020 .

Head of department _____ A.B. Kassymova, assoc. professor, PhD
signature full name, title, degree

Author _____ A.A. Kuatbayeva, assistant-professor, Ph.D. in CS
signature full name, title, degree

The educational-methodological complex of discipline was approved at the meeting of the Academic Council of JSC "International information technologies university" Minutes №____1____ dated "_28"___08____2020 .

Director of the Department _____ A.Mustafina
Signature

for Academic Affairs

1. General information	
Faculty	Information Technology
Major code and title	5B070300 Information Systems
Year, semester	3 year, 5
Subject category	Compulsory Elective Profiling
Number of Credits	5
Language of Delivery:	English
Prerequisites:	OOP, Probability theory and mathematical statistics
Postrequisites	Business analytics
Lecturer	Kuatbayeva A.A., Ph.D. in CS, assistant- professor,802, Tel.:+77089731932 Email: Ahamala2017@gmail.com, ahamala2020@gmail.com, a.kuatbayeva@edu.iitu.kz MS teams – a.kuatbayeva consultation schedule: 10.00-14.00 Wednesday

2. Goals, objectives and learning outcomes of the course

The course goal is

training the specialists, capable to use SAS programming technologies for the solution business data analysis tasks. Description the basic SAS/BASE programming, macros concepts, SQL, SAS/STAT statistical analysis methods library: dispersion analysis, linear regression, logistic regression, the generalized linear models, Tobit models, the survival analysis, clustering. Bayesian analysis; interactive programming language for processing IML matrix data: matrix, linear algebra and numerical methods in IML, integration with SAS/BASE, R, C ++; explanation SAS/BASE programming technology aspects (syntax, semantics, rules, ways to work with it).

Course objectives

Learn how to generate descriptive statistics and explore data with analysis of variance graphs perform and apply multiple comparison techniques, perform linear regression and assess the assumptions, use regression model selection techniques to aid in the predictor variables choice in multiple regression, use diagnostic statistics to assess statistical assumptions and identify potential outliers in multiple regression, use chi-square statistics to detect associations among categorical variables, fit a multiple logistic regression model.

Lectures are supported by practical tasks, performed by listeners independently on SAS OnDemand for Academics technology. By the performed practical tasks and examination results, the students who have fulfilled successfully course requirements besides a positive assessment for a special course will receive the SAS company's certificate.

Learning Outcomes:

Know the basic standards, technologies, and notations of business intelligence modeling; Tools used by business intelligence The main applications of modeling in Business Intelligence.

To be able to model, analyze and improve business intelligence algorithms using learned standards, technologies; to review business analytics model;

Generate business intelligence code.

Be competent in using SAS-based business intelligence methods and tools.

3. Course description

Demonstrate knowledge and comprehension on SAS programming basic ideas and concepts such as macros, SQL, SAS/STAT statistical analysis methods library: dispersion analysis, linear regression, logistic regression, the generalized linear models, Tobit models, the survival analysis, clustering. Bayesian analysis; interactive programming language for processing IML matrix data: matrix, linear algebra and numerical methods in IML; integration with SAS/BASE, C++; explanation SAS/BASE programming technology aspects (syntax, semantics, rules, ways to work with it).

Students also perform independent work under the guidance of a teacher (IWST) and independent work of a master (IWS) on certain topics.

4. Course policy

Students are forbidden to:

- submit any tasks after the deadline. The mark for late submissions is decreased;
- cheat. Plagiarized papers shall not be graded and receive a "0";

- be late for classes;
- retake any tests, unless there is a valid reason for missing them;
- use mobile phones in class;

Students should always

- be appropriately dressed (formal/semi- formal styles are acceptable);
- show consideration for and mutual support of teachers and other students;
- let the teacher know of any problems arising in connection with their studies.

During classes students can use online platforms like MS Teams, DL, Zoom, etc.

5. Literature

Basic literature:

Literature (with an indication of the authors and data output), the availability (number), software and consumables with information about where you can get them. (1-6)

1. SAS/STAT(R) 9.4 User's Guide
2. SAS(R) 9.4 Functions and CALL Routines: Reference Base
3. SAS(R) 9.4 Procedures Guide
4. SAS(R) 9.4 SQL Procedure User's Guide
5. SAS (R) 9.4 Data sets options Reference
6. SAS (R) 9.4 Macro language Reference
7. Video course of video lectures and practical classes in dl.iitu.kz- 11728 SAS Programming 2020-2021/1 – Video course folder in materials- https://dl.iitu.kz/mod/folder/view.php?id=215030
8. https://support.sas.com/edu/elearning.html?ctry=us&productType=library

3. **Online Resources & References**

Video course of video lectures or reference to it and practical classes in dl
https://dl.iitu.kz/mod/folder/view.php?id=215030
MS Teams,
https://support.sas.com/edu/elearning.html?ctry=us&productType=library
Academic program CIS: https://www.sas.com/ru_ru/academic/overview.html
Documentation: http://support.sas.com/documentation/
Software: SAS University Edition http://www.sas.com/en_us/software/university-edition.html

6. Course schedule

Week/ date	Course topics	References	Lectures (h/w)	Practical sessions	Lab. sessions	TSIS (h/w)	SIS (h/w)
1	Lecture 1.Introduction. SAS analytic platform, technologies review. SAS U edition. SAS Data sets. SAS libraries. Format and options for Data sets in SAS.	[1], [2]	1		2	6	
2	Lecture 2. SAS/BASE Programming language bases. SAS libraries. Variables attributes in SAS. Date and time. DATA step. Conditional cycles in DATA step. Arrays in DATA step, their features.	[1],[2], [4], [6]	1		2	6	
3	Lecture 3. Data sets options: keep, drop, rename, where. DATA Step features: a) two SET operators; b)two data sets in one SET (append/union). The running sums. Data sets sorts. Group mining in DATA step. Splitting data sets. Queue for variable. Proc Transpose.	[1],[2], [4], [6]	1		2	6	
4	Lecture 4. SAS Macro. Macrosubstitutions. Macro variables. Macros run order. Debugging macros. Parameters in macros, macro functions. Cycles.	[2], [3]	1		2	6	
5	Lecture 5. SQL usage and implementation features in SAS. PROC SQL main operators. Access to other DBMS. PROC SQL and macro language.	[2], [3]	1		2	6	

6	Lecture 6. SQL & Reporting. SAS formats and their usage. Reports creation on DATA step. Reporting via PROC SQL, PROC PRINT, PROC TABULATE. Tool selection depending on report type. Output delivery system overview. SAS University Edition features.	[2]	1		2	6	
7	Lecture 7. SAS/STAT statistical analysis methods library (MODULE I): dispersion analysis. Examining data distributions, obtaining and interpreting sample statistics using the UNIVARIATE and MEANS procedures, examining data distributions graphically in the UNIVARIATE and SGPLOT	[1], [2]	1		2	6	
8	*Midterm control*	[1], [2], [3], [4], [5], [6]	1		2	6	
9	Lecture 8 SAS/STAT statistical analysis methods library Lecture 9. SAS/STAT statistical analysis methods library (MODULE I): Logistic regression.	[1], [2], [4], [6]	1		2	10	
10	Lecture 10. SAS/STAT statistical analysis methods library (MODULE II): Analysis of Variance (ANOVA). Target Marketing. Churn Prediction. Credit Scoring, Collection Scoring. Fraud Detection.	[1-6]	1		2	8	
11	Lecture 11. SAS/STAT statistical analysis methods library (MODULE II): The generalized linear models (Poisson regression, negative binomial regression, etc.)	[1-6]	1		2	7	
12	Lecture 12. SAS/STAT statistical analysis methods library (MODULE III): Tobit models in SAS/STAT	[1-6]	1		2	10	
13	Lecture 13. SAS/STAT statistical analysis methods library (MODULE III): The SAS/STAT survival analysis.	[1-6]	1		2	8	
14	Lecture 14. Interactive programming language for processing IML matrix data. Introduction to IML. Integration with SAS/BASE, C++. SAS/IML SAS/IML Linear algebra and Numerical methods. Practical cases, working on projects.	[1-6]	1		2	7	
15	*Endterm*	[1-6]	1		2	7	
	Total hours:	**150**	**15**		**30**	**105**	

6. List of topics/ assignments for laboratory classes

№	Topic Title	Number of hours	References	Form of reporting	Deadline
1	2	3	4	5	6
1	SAS U edition overview.	2	[1-6]	Code on SAS Base	2
2	Writing simple programs on SAS/BASE assigned to DATA step conditional cycles and work with arrays.	2	[1-6]	Code on SAS Base	3
3	Work with data sets.	2	[1-6]	Code on SAS Base	4

4	Programming with SAS Macro.	2	[1-6]	Code on SAS Base	5
5	Individual tasks for programming on SAS Macro	2	[1-6]	Code on SAS Base	6
6,7	Work with SQL and reporting. Programming tasks with SAS/STAT library (MODULE I) on dispersive analysis.	4	[1-6]	Code on SAS Base	7,8
8	Programming tasks with SAS/STAT library for linear regression (MODULE I). Linear Regression Diagnostics: examining residuals; investigating influential observations; assessing collinearity.	2	[1-6]	Code on SAS Base	9
9	Work on logistic regression tasks	2	[1-6]	Code on SAS Base	10
10, 11	Work on ANOVA tasks. Work with generalized linear models (Poisson regression, negative binomial regression, etc.)	4	[1-6]	Code on SAS Base	11,12
12	Work with Tobit models in SAS.	2	[1-6]	Code on SAS Base	13
13-15	Work with survival analysis tasks. Programming for IML:Matrixes, linear algebra, numerical methods, integration with SAS/BASE, C++. Work with linear algebra and numerical methods in IML. Work in a groups on cases.	6	[1-6]	Code on SAS Base	14, 15

7. List of topics/assignments for Student Independent Study

Proper organization of students independent study is the key to the formation of skills in mastering, learning, assimilation and systematization of acquired knowledge, ensuring a high level of academic performance in the learning process

№	Topic/Assignment title	Number of hours	References	Form of reporting	Deadline
1	2	3	4	5	6
1	Programming on SAS/BASE according tasks.	13	[2], [1]	Write code in a SAS Base	3
2	Individual tasks for programming on SAS Macro	13	[3], [4]	Write code in a SAS Base	5
3	Individual tasks for programming on SAS/STAT (MODULE I)	13	[3], [4], [6]	Write code in a SAS Base	7
4	Work on logistic regression tasks	13	[2], [5]	Write code in a SAS Base	9
5	SAS/IML Linear algebra and Numerical methods	13	[1], [3], [4]	Write code in a SAS Base	11
6	Individual tasks for programming on SAS/IML.	13	[3], [5], [6]	Write code in a SAS Base	13
7	Tobit models	12	[2], [1]	Write code in a SAS Base	14

10. System for evaluating student performance in a discipline:

Option 1

Period	Assignments	Score	Total
1st attestation	**Laboratory works:** Lab work 1, Lab work 2, Lab work 3, Lab work 4, Lab work 5, Lab work 6, Lab work 7, **Team project +certificates** **Mid-term** **SIS assignments**	**35** **5** **5** **5** **5** **5** **5** **25** **25** **15**	**100**
2nd attestation	**Laboratory works:** Lab work 8, Lab work 9, Lab work 10, Lab work 11, Lab work 12, Lab work 13, Lab work 14, Lab work 15, **Team project +certificates** **Mid-term** **SIS assignments**	**35** **3** **4** **5** **5** **5** **5** **5** **3** **25** **25** **15**	**100**
Exam			**100**
Total	**0,3*1stAtt+0,3*2ndAtt+0,4*Final**		

*If the number of absences exceeds 20%, student will be automatically scheduled for a Retake (summer semester)

Option 2

Each type of educational work is evaluated on a 100-point scale and is included in the average assessment of the current control, taking into account the weighting coefficient in accordance with the table

Period	Assignments	Maximum score	Weighting coefficient	Total
1st attestation	Laboratory practice	100	0,2	**100**
	Course project	100	0,3	
	IWS	100	0,3	
	Attending lectures	100	0,1	
	Working for certificates achievement	100	0,1	
2nd attestation	Laboratory practice	100	0,2	**100**
	Course paper	100	0,3	
	IWS	100	0,3	
	Attending lectures	100	0,1	
	Working for certificates achievement	100	0,1	
Exam				**100**
Total	**0,3*1stAtt+0,3*2ndAtt+0,4*Ex**			**100**

*If the number of absences exceeds 20%, student will be automatically scheduled for a Retake (summer semester)

11. Assessment criteria:

Option 1

Example of assessment criteria on a 5-point scale for laboratory works:

Points	Assessment criterion
5	The work was completed in full and correct answers were received for additional questions from the teacher within the framework of the program.
4	The work was completed in full, but mistakes were made when answering additional questions from the teacher.
3	The work was completed in full, correct conclusions were made, however, there are some non-compliance with design requirements, for example, errors in the design of graphs, tables, or in recording measurement results. After teacher's instructions, these shortcomings are eliminated.
2	The work was performed in an incomplete volume, for example, fault calculations were not carried out or carried out incorrectly, some results are incorrect, the conclusions do not correspond to reality, there are significant errors in the graphical data. After teacher's instructions, the main shortcomings were eliminated, and the graphs were corrected.
1	Work is performed in an incomplete volume, for example, there are errors in the calculations of most or all of the desired values, no faults, the results are mostly present, but not true, the conclusions do not correspond to reality, there are significant errors in the design, there are no graphs, calculation formulas are not specified, etc. After teacher's instructions, the main shortcomings are eliminated.
0	The work is incomplete, for example, there are errors in the calculations of most or all of the required values, there are no faults, the results are mostly present, but not true, the conclusions do not correspond to reality, there are significant errors in the design, there are no graphs, calculation formulas are not specified, etc.

Option 2

The point-rating letter system for assessing the educational achievements of students with their interpretation in the traditional grading scale:

Letter Grade	Numerical equivalent	Points (%)	Traditional system assessment	General description of grading criteria
A	4,0	95-100	Excellent	The student has knowledge of the subject in the full scope of the curriculum, understands the discipline deeply enough; shows a high level of knowledge that exceeds the volume provided by the syllabus, gives an exhaustive answer
A-	3,67	90-94		The student has knowledge of the subject in the full scope of the curriculum, understands the discipline deeply enough; gives an exhaustive answer
B+	3,33	85-89	Good	The student shows a complete, well-founded knowledge of the subject, but the answers did not always highlight the main idea, rational methods of calculation were not always used; the answers were mostly brief and sometimes unclear.
B	3,0	80-84		
B-	2,67	75-79		
C+	2,33	70-74		

C	2,0	65-69	Satisfactory	The student demonstrates sufficient knowledge of the subject, but without proper depth and justification, the answers are unclear and without proper logical sequence.
C-	1,67	60-64		
D+	1,33	55-59		
D	1,0	50-54		
FX	0,5	25-49	Unsatisfactory	The student demonstrates insufficient knowledge of the subject, positive answers were not given to individual questions.
F	0	0-24		The student demonstrates a very low level of knowledge of the subject.

8. Assessment and evaluation materials (exam questions)

Project

ECG analysis is an important part of finding devices, especially in the cardiovascular system. Various parameters obtained with the ECG curve, such as P-R, QT, R-R, Q-R-S, and S-T intervals, as well as heart rate variability, serve as safety and efficacy endpoints for many clinical trials. There are different algorithms based on graphs and frequency domains to detect different waveforms and ECG cells [4]. This article discusses the study and determination of ECG using SAS. ECG data, which are used for analysis and articles in clinical, are usually processed. The study of the algorithm of ECG signs using SAS allows you to flexibly and clearly perform a preliminary analysis of the parameters obtained from the ECG. In addition, it allows the programmer or researcher to analyze the same data from a research point of view.

ECG signals data mining

ISSUE

Given the dataset with ECG signals (4000 rows). Using SAS UE you need to determine the algorithms on SAS for it.