

Course Syllabus

Title of the course	Databases (offered in English)			
Title of the Academic Programme	BA “Sociology and Social Informatics”			
Type of the course	Elective			
Prerequisites	Information Systems, Algebra and Analysis			
ECTS workload	5			
Total indicative study hours	Directed Study	Self-directed study	Total	
	44	146	190	
Course Overview	<p>This course introduces database design and querying in different Database Management Systems. Emphasis is both on database design and on applications of databases to analytical tasks.</p> <p>We will cover database theory, characteristics of contemporary DBMS landscape, Relational Algebra, Entity Relationship modeling of different domains, SQL and its dialects, different approaches to data modelling.</p> <p>Course DB project will include group modelling of a complex domain area, design, and implementation of DB for this area, including relationships, triggers, stored procedures, complex queries.</p> <p>Applications of databases to Business Intelligence are also discussed through the course, including the integration of predictive analytics with DBMS and user BI systems.</p> <p>Course BI project will include defining BI goals, setting KPI and designing dashboards for a particular case from different domain areas. Visualization of business metrics and dashboards design will become a special topic of the course, which will allow students to get familiar with the process of data communication. Data-driven dashboards, being one of the most essential parts of Business Intelligence, will be discussed a lot.</p>			
Intended Learning Outcomes (ILO)	<p>Upon completion, students should be able to:</p> <ul style="list-style-type: none"> - model different domain areas using ER approach - design and implement normalized database structures by creating database tables, queries, triggers, stored procedures, reports, and forms - perform typical BI reporting queries using SQL and analytical tools - understand applications of BI to decision support in modern companies - produce custom reports and dashboards based on DB data in Tableau and/or Power BI 			
Teaching and Learning Methods	Teaching and learning methods include lectures, tutorials, project work, home assignments.			
Content and Structure of the Course				
№	Topic / Course Chapter	Total	Directed Study	Self-directed

			Lectures	Tutorials	Study
1	Fundamentals of Databases: Relational Algebra, ER modelling	40	0	6	34
2	Database Design	26	2	4	20
3	SQL	80	0	20	60
4	Dashboards design and basics of BI	44	2	10	32
Total study hours		190	4	40	146
Indicative Assessment Methods and Strategy	<p>Assessment strategy is designed to cover both practical skills and teamwork via group projects and theoretical knowledge via 3 tests. The overall grade consists of:</p> <ul style="list-style-type: none"> • Seminar participation (10%) • In-class test on ER diagrams and Relational Schemas (10%) • In-class test on SQL queries (10%) • DB Project (25%) • BI Project (25%) • 80-minute final test (20%) <p>Seminar participation can be replaced with extra homework.</p> <p>If the student were not able to attend the in-class test without a valid excuse, the in-class test can be retaken in the next two days with its score reduced by 1 point.</p> <p>If the student were not able to attend the in-class test <u>with</u> a valid excuse, the in-class test can be retaken without any penalties. Confirmation of the valid excuse is required. Retake time is negotiated individually.</p> <p>If the project was submitted an hour after the deadline, the score for it is reduced by 1 point, 6 hours - by 5 points, after 7 hours after the deadline projects are not accepted.</p> <p>The final test cannot be retaken.</p> <p>In case of failing the course, it could be retaken during the additional examination period by passing some parts of DB (several SQL queries) and BI projects (dashboard) and 10 questions from the final test.</p> <p>Grades are rounded only once at the end of the course. The roundup follows the general math rules for rounding, where the number to the right of the rounding digit determines whether the number is rounded upwards or downwards.</p>				
Readings / Indicative Learning Resources	<p><u>Mandatory</u></p> <p>Clare Churcher. <i>Beginning Database Design: From Novice to Professional</i>. 2 edition. New York, NY: Apress, 2012. https://link.springer.com/book/10.1007%2F978-1-4302-4210-9</p> <p>Foster, Elvis C., and Shripad V. Godbole. <i>Database Systems</i>. Berkeley, CA: Apress, 2014. https://library.books24x7.com/toc.aspx?bookid=78479</p> <p>Stanford's Introduction to Databases MOOC https://lagunita.stanford.edu/courses/Engineering/db/2014_1/ab out</p> <p><u>Optional</u></p> <p>Khan, Arshad. <i>Jumpstart Tableau: A Step-By-Step Guide to Better Data Visualization</i>. Apress, 2016. https://library.books24x7.com/toc.aspx?bookid=115155</p>				

	<p>Laursen, Gert H. N., and Jesper Thorlund. <i>Business Analytics for Managers: Taking Business Intelligence Beyond Reporting</i>. 1 edition. Hoboken, N.J: Wiley, 2010. https://library.books24x7.com/toc.aspx?bookid=43172</p> <p>Monsey, M. and Sochan, P. <i>Tableau For Dummies</i>. John Wiley & Sons, 2016. https://library.books24x7.com/toc.aspx?bookid=105577</p> <p>Rockoff, L. <i>The Language of SQL</i>. Course Technology Cengage Learning, 2011. https://library.books24x7.com/toc.aspx?bookid=42732</p>		
Indicative Self- Study Strategies	Type	+/-	Hours
	Reading for seminars / tutorials (lecture materials, mandatory and optional resources)	+	26
	Assignments for seminars / tutorials / labs	+	60
	E-learning / distance learning (MOOC / LMS)	-	
	Fieldwork	-	
	Project work	+	60
	Other (please specify)	-	
	Preparation for the exam	-	
Academic Support for the Course	Academic support for the course is provided via email		
Facilities, Equipment and Software	Computer class, PostgreSQL/MSSQL Server, Tableau and/or PowerBI, diagramming software, MSSQL Server Studio, MS Access		
Course Instructor	Anastasiya Kuznetsova		

Intended Learning Outcomes (ILO) Delivering

Programme ILO(s)	Course ILO(s)	Teaching and Learning Methods for delivering ILO(s)	Indicative Assessment Methods of Delivered ILO(s)
UC-3 - Ability to solve problems in professional practice based on analysis and synthesis	The student is able to deal with real cases managing tasks and solving problems related to the basic analysis in the field of DBMS.	Lectures, Tutorials, Home assignments	In-class tests, seminar participation
UC-5 - Ability to work with information: find, evaluate and use necessary information from different sources in order to solve	The student is able to work with the information from different sources for solving scientific	Tutorials, Home assignments	DB and BI projects

scientific and professional problems (including through the system approach)	and professional problems as well as data extraction with the use of SQL syntax.		
UC-7 - Ability to work as part of a team	The student is able to organize work in group projects.	Home assignments, Self-study	DB and BI projects
PC-5 - Ability to use the basic laws of natural sciences in professional activities, apply the methods of mathematical analysis and modeling and theoretical and experimental research when solving professional problems	The student is able to apply ER-diagrams and Relational schemas to real-world problems.	Tutorials	In-class tests, seminar participation
PC-10 - Ability to process and analyze data to prepare analytical decisions, expert advice, and recommendations	The student is able to design, create databases and BI dashboards with the use of SQL syntax.	Tutorials, home assignments	DB and BI projects, seminar participation
PC-11 - Ability to plan and carry out project works in the fields of public opinion research and organization of work of marketing services providers	The student is able to plan, carry and organize group project and produce custom reports and dashboards based on DB data in Tableau and/or Power BI.	Tutorials	DB and BI projects
PC-12 - Ability to participate in analytical and consulting activities	The student is able to understand applications of BI to decision support in modern companies and participate in analytical and consulting activities.	Lectures	DB and BI projects, seminar participation

PC-13 - Ability to use the methods of collecting, processing, and interpreting comprehensive social information to solve organizational and managerial tasks, including those outside the immediate scope of activities	The student is able to perform typical BI reporting queries using SQL and analytical tools.	Tutorials, Home assignments	DB and BI projects
---	---	-----------------------------	--------------------

Course Content

1. Fundamentals of Databases: Relational Algebra, ER modelling

1.1 *Theoretical introduction to databases*

Relational and non-relational databases. SQL databases. Introduction to set theory and relational algebra.

1.2 *Entity-relationship diagrams and Relational schemas*

Usage of ER diagrams. ER diagram symbols. Relational schemas design. Keys in relational schemas and main terminology.

2. Database Design

2.1 *Databases design*

Designing databases in Lucidcharts, MySQL and draw.io.

2.2 *Database creation*

Connection to databases. Creation of databases with diagrams and code.

3. SQL

3.1 *Introduction to SQL queries*

Practice with functions SELECT (*), WHERE, AND/OR.

3.2 *Queries with constraints*

Practice with functions: SELECT, ORDER BY, LIKE, IN, LIMIT.

3.3 *Queries with several conditions. Filtering and sorting*

Practice with functions: WHERE, ORDER BY, NULL.

3.4 *Introduction to dataset joins*

Types of joins and their usage. Practice with JOIN function.

3.5 *Data aliases*

Practice with joins and AS function.

3.6 *Aggregate functions in SQL*

Practice with functions: MIN, MAX, AVG, COUNT, GROUP BY, HAVING.

3.7 *Complex SQL queries*

SQL subqueries for data preprocessing.

3.8 *Data modification*

Practice with functions: INSERT, UPDATE, DELETE, CREATE / TEST.

4. Dashboards design and basics of BI

4.1 *Dashboard design*

The history of the dashboards, getting familiar with what KPI metrics are. Types of information dashboards. Main rules for making dashboards and common mistakes.

4.2 Introduction to Tableau

Types of data connections in Tableau. How to connect SQL database to Tableau. Basic diagrams in Tableau (bar charts, scatterplots, line graphs, pie charts, tables).

4.3 SQL and Tableau

SQL queries in Tableau for getting better data for your dashboards. Data joins and filtering with queries.

Assessment Methods and Criteria

Assessment Methods

Types of Assessment	Forms of Assessment	Modules			
		1	2	3	4
Formative Assessment	Seminar participation	*	*		
	Test	*	*		
	Project		*		
	Final test		*		

Assessment Criteria

Seminar participation

Grades	Assessment Criteria
«Excellent» (8-10)	A critical analysis which demonstrates original thinking and shows strong evidence of preparatory research and broad background knowledge. Complete all tasks during the seminars and practice lessons.
«Good» (6-7)	Shows strong evidence of preparatory research and broad background knowledge. Excellent oral expression. Have some problems with seminar tasks but show the whole understanding of the subject.
«Satisfactory» (4-5)	Satisfactory overall, showing a fair knowledge of the topic, a reasonable standard of expression. Some hesitation in answering follow-up questions and/or gives incomplete or partly irrelevant answers. Complete only half of the tasks at the seminars.
«Fail» (0-3)	Limited evidence of relevant knowledge and an attempt to address the topic. Unable to offer relevant information or opinion in answer to follow-up questions. Does not complete seminar tasks.

Project Work

Example of the DB project

Choose the sphere in which you are interested, design database for it and make a useful report using database queries. Create a database with connected tables (by queries or visually) taking into account correct names of columns, data types, etc. Insert values.

Make a report which consists of queries (10-15) being guided by “one conclusion = one query” principle. Queries should include as many functions and keywords that have been covered at classes as you can. Every part (query) of the report should have practical meaning (conclusions that could be used). You can work in pairs and send final presentation (all parts with explanations, database structure, inserts) via email.

Example of the BI project

In this homework assignment, you have to make a dashboard based on the questions below. You have to make different diagrams for each question and create a dashboard based on it. Publish your final dashboard to Tableau Public and send the link to that. You can work in pairs and send the final dashboard via email.

1) Where are our customers from? Create new categories of regions (North - AC, AP, AM, PA, RO, RR, TO; Northeast - AL, BA, CE, MA, PE, PI, RN, SE; Central-West - GO, MT, MS, DF; Southeast - ES, MG, RJ, SP; South - PR, RS, SC)

You can make a map but don't forget to edit the country at “Edit Locations” to Brazil. Also, add the filter with these new categories.

2) What are the TOP 15 states, from which customers make orders which cost (payment value) more than the average?

3) What are the most popular types of payments? (Show TOP 10).

4) What is the most frequent month and weekday when our customers order items (take order approved date)?

5) Is there any relation between payment value and a number of instalments?

6) What is the average payment value and average delivery time (the difference between approved and delivered to the customer dates)? (Show as a table.)

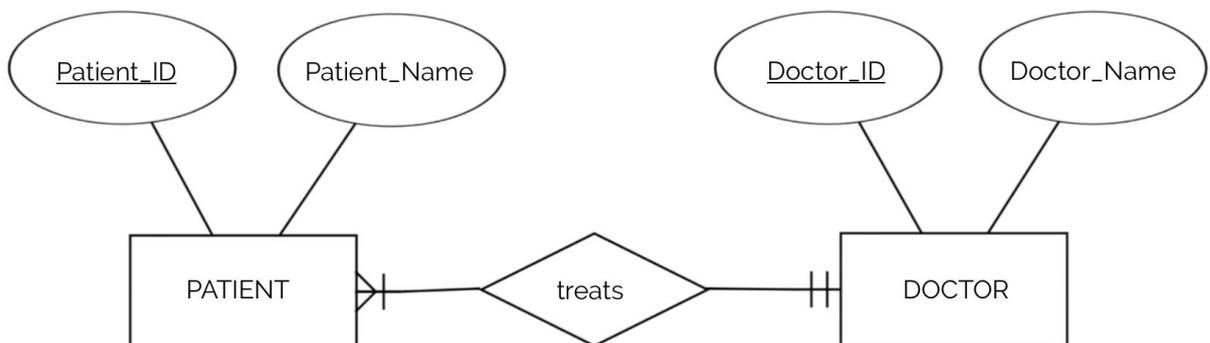
Grades	Assessment Criteria
«Excellent» (8-10)	<p>Shows strong evidence and broad background knowledge. In a group presentation all members contribute equally and each contribution builds on the previous one clearly; Answers to follow-up questions reveal a good range and depth of knowledge beyond that covered in the presentation and show confidence in discussion.</p> <p>For DB project: all queries were correct (only two mistakes are possible). There is a demonstration of understanding and using the most covered topics during the course. The project is organized clearly and inferred logically.</p> <p>For BI project: all questions were answered correctly, diagrams correspond to the data. Main principles of dashboards design are observed. Filters and calculated values work correctly.</p>
«Good» (6-7)	<p>Clearly organized analysis, showing evidence of a good overall knowledge of the topic. The presenter of the project work highlights key points and responds to follow up questions appropriately.</p> <p>For DB project: most queries were done correctly, but there are some problems with their logic and justification of conclusions.</p> <p>For BI project: 70% of question were asked correctly, diagrams correspond to the data, dashboard design principles were followed, have some problems with filters and calculated fields.</p>
«Satisfactory» (4-5)	<p>Takes a very basic approach to the topic, using broadly appropriate material but lacking focus. The presentation of project work is largely unstructured, and some points are irrelevant to the topic. Knowledge of the topic is limited and there may be evidence of basic misunderstanding.</p>

	<p>For DB project: the majority of queries are incorrect and shows poor student's knowledge acquired during the course. Understanding of basic principles of working with databases and writing queries are not presented.</p> <p>For BI project: half of the questions were answered correctly, dashboard design principles were not followed. No filters and calculated fields. Diagrams do not correspond to the data.</p>
«Fail» (0-3)	<p>Fails to demonstrate any appropriate knowledge.</p> <p>For DB project: almost no queries are correct, the main tasks were failed and there is no understanding of basic principles of working with.</p> <p>For BI project: less than a half of questions were answered correctly, dashboard design principles were not followed. No filters and calculated fields. Diagrams do not correspond to the data.</p>

Written Assignments (Test)

Example of the questions in the test on ER diagrams and relational schemas

1. Make an ER diagram that describes the following data and relationships:
 - An organization keeps track of its customers and their favorite shops.
 - For each customer it keeps track of the CustomerID (unique), CustomerName, and CustomerDB (date of births).
 - For each shop it keeps track of the ShopID (unique) and ShopName.
 - Each customer has only one favorite shop and each shop must have at least one customer who likes this shop, but can have many customers.
2. Make an ER diagram that describes the same data and relationships as in the 1st task but modify this:
 - Each customer has one and many favorite shops, and each shop location must have at least one customer who likes this shop but can have many customers.
3. Make a relational schema that represents data and relationships from this diagram:



4. Describe in the words the data that would be contained in this table?

VEHICLE

VecihleNumber

State Brand Model Year

Example of the questions in the test on SQL queries

1. Find out the number of all women who have white or asian race, who are older than 30 but younger than 33 and have a pay rate more than or equal to 20. Write your query below.
2. In table core dataset find out employee name of all people that have such states: ND, MT, OR, AZ, ME, RI, VT, PA, MA.
3. Write a query to retrieve all the distinct customer_names with their cities and favourite shops.
4. Find out the information below about the stations that necessarily had lines and were closed. Order results by name of the line and name of the station (a lot of 9 - the station is open). Name the first station.

Example of the questions in the final tests

This test includes two parts. The first one is based on our lectionary and theory which we had during the course. The second one is based on SQL practice - you have to make SQL queries with data provided and answer the questions about this data.

1. 2pt Which standard aggregation functions you can use in SELECT statements?
(a) count
(b) first
(c) greatest
(d) avg
(e) last
2. 1pt What statements are used to modify records in tables?
(a) modify
(b) change
(c) update
(d) convert
3. 3pt Write at least 3 rules for good dashboard design.
4. 2pt Insert the missing keywords and object names of the database for the SQL statement, which add a string (114, 'Mike', 18).
Template:
(a) INTO PERSON(ID, NAME, AGE) (b) (114, 'Mike', 18);
5. How do we call columns which help us to join data tables with the analysis of Relational schemas?

6. How many products were sold in January 2015?
7. What is the average amount of profit made by the company per day?
8. What country has the third place by the number of products sold?
9. What is the most frequent weekday when our customers order items?
10. In what city and state is the store that had the greatest total sum of sales?

Grades	Assessment Criteria
«Excellent» (8-10)	75% of points earned or more
«Good» (6-7)	50% - 74% of points earned
«Satisfactory» (4-5)	30% - 49% of points earned
«Fail» (0-3)	Less than 30% of points earned

Recommendations for students about organization of self-study

Self-study is organized in order to:

- Systemize theoretical knowledge received at lectures;
- Extending theoretical knowledge;
- Learn how to use legal, regulatory, referential information and professional literature;
- Development of cognitive and soft skills: creativity and self-sufficiency;
- Enhancing critical thinking and personal development skills;
- Development of research skills;
- Obtaining skills of efficient independent professional activities.

Self-study, which is not included into a course syllabus, but aimed at extending knowledge about the subject, is up to the student's own initiative. A teacher recommends relevant resources for self-study, defines relevant methods for self-study and demonstrates students' past experiences. Tasks for self-study and its content can vary depending on individual characteristics of a student. Self-study can be arranged individually or in groups both offline and online depending on the objectives, topics and difficulty degree. Assessment of self-study is made in the framework of teaching load for seminars or tests.

In order to show the outcomes of self-study it is recommended:

- Make a plan for 3-5 presentation which will include topic, how the self-study was organized, main conclusions and suggestions and its rationale and importance.
- Supply the presentation with illustrations. It should be defined by an actual task of the teacher.

Special conditions for organization of learning process for students with special needs

The following types of comprehension of learning information (including e-learning and distance learning) can be offered to students with disabilities (by their written request) in accordance with their individual psychophysical characteristics:

1. *for persons with vision disorders:* a printed text in enlarged font; an electronic document; audios (transferring of learning materials into the audio); an individual advising with an assistance of a sign language interpreter; individual assignments and advising.
2. *for persons with hearing disorders:* a printed text; an electronic document; video materials with subtitles; an individual advising with an assistance of a sign language interpreter; individual assignments and advising.
3. *for persons with muscle-skeleton disorders:* a printed text; an electronic document; audios; individual assignments and advising.