RESIT COURSEWORK HANDOUT

BUSINESS INTELLIGENCE

MODULE CODE	CCP6415
MODULE TITLE	Business Intelligence
PROGRAMME	MSc in Advanced Software Engineering - Big Data Engineering and Data Science
DEPARTMENT	Computer Science
CREDITS	15
STAGE OF STUDY	Postgraduate
SEMESTER/SESSION	Autumn 2024-2025
LOCATION	Athens
STAFF	K. Koutsoulelos
E-MAIL	kkoutsoulelos@athtech.gr
STAFF OFFICE	Sofias Sliman 3

ASSESSMENT NUMBER	Resit Coursework
ASSESSMENT TITLE	Real life project
ASSESSMENT TYPE	Individual Assignment
HAND-OUT DATE	17/07/2025
SUBMISSION DATE	01/09/2025
FEEDBACK DATE	10/09/2025

LEARNING OUTCOMES

Upon completion of this piece of assessment, a student will be able to:

- LO1: analyse business intelligence key concepts and its relevant theory and techniques,
- LO2: outline the major frameworks of computerised decision support systems,
- LO3: breakdown different aspects of building, deploying and managing BI systems,
- LO4: identify and examine business and technical requirements for a BI solution,
- LO5: describe the opportunity for business intelligence and efficient performance management in business,
- LO6: distinguish and comfortably manipulate strategic parameters and important issues of BI,
- LO7: value the importance of databases, and data warehouses,
- LO8:adequately implement and support topics regarding data integration and data visualisation
- LO9: efficiently and accurately interpret BI related application areas,
- LO10: use Python for BI purposes.

ASSESSMENT CRITERIA

Reports and Project

- Task completeness
- Application of theoretical framework and delivered techniques
- BI solution development status
- Justification of decisions and recommendations
- Reference and investigation of related industry
- Report structure and presentation

DETAILED DESCRIPTION

Scenario

This integrative group project is designed to encourage students to apply the knowledge and experience learned in class towards a real-life business intelligence system.

In this context, students will be in need to address the requirements of a client which is a veterinary clinic and sufficiently cover the below mentioned points and tasks.

The final desired outcome is to help veterinarians get insights on what are the possible effects of a prescription.

The Company

The client is a veterinary clinic that employs experienced veterinarians, and they want to assist them in the decision-making process of what drugs/prescriptions should be avoided for the animals they are looking after.

They need a centralised view of historical data of adverse events reported in FDA for the last 15 years in order to help them answer faster and more accurate questions including but not limited to:

- What are the most common reactions for every breed
- What are the most common active ingredients that cause side effects
- How the size of an animal (height/weight) is correlated with the reactions and the outcome of the event
- How the gender of an animal and its reproductive status is correlated with the reactions and the outcome of the event
- What is the geographic distribution of those events
- How many days it takes for the reactions to appear

In addition, specialised in dogs veterinarians would also like to check this information on a higher level when possible, meaning that they want to have overall statistics for:

- Breeding Groups (Herding, Hound, Toy e.g)
- Breeding Purpose (Guarding, Companion e.g)

Note that for several reasons, such as security or privacy, the drug names and manufactures are not always provided and you may see MSK (masked null) as value for them. The implementation and the analytics should be based on the active ingredients (for example Drug: Depon Active Ingredient: Paracetamol) where it is appropriate.

Existing Data Sources

FDA provides the data of all the adverse events reported to them in a detailed way. You can either extract the data from the API or download the files via an automated process:

- API: https://open.fda.gov/apis/animalandveterinary/event/
- Files: https://open.fda.gov/data/downloads/

For the breeding information you can utilise one of the free APIs that exists for dogs and their breeds (https://api.thedogapi.com/v1/breeds for instance).
Note: It is expected to have at least one more extra source for breeds (any species) except the one mentioned above. This source can be anything like a public API, files, data dump etc.

You may find complementary data for the reactions in:
https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/combined-veterinary-dictionary-drug-regulatory-activities-veddra-list-clinical-terms-reporting_en.pdf

Tasks

1. Information Visualization & Data Analytics **[70%]**
    a. Business Reports (types, usage, target areas/ users)
    b. Data/ Information Visualization and Dashboard design and description of functionality and intended purposes

2. Conclusion **[30%]**
    a. Synopsis of the new proposed BI and decision support system
    b. Summary of the BI solution's advantages and limitations
    c. Reference to implementation issues, new achieved workflows and change management issues
    d. Proposals for further improvements and future steps

Note: The report should open with an introduction covering the team's tasks of the coursework. At a minimum should include:
- Schema design rationale (key entities, relationships, normalization/denormalization choices)
- Data-point selection & elimination (cleaning steps, filtering criteria, outlier removal)
- Initial data preparation (transformation rules, enrichment sources)

SUBMISSION
Tasks should be solved individually. Each student has to separately submit a report and other supporting files (git repositories or code files/scripts) concerning this task. The format of the reports should adhere to the following characteristics: the font should be Times New Roman 12pt, line spacing should be single. Students are expected to submit their work on TURNITIN. Late submissions are subject to a 5% penalty for every day of delay. Any reports submitted after 7 days from the submission date will get a mark of 0 (zero). The interim presentation must be submitted via email before the beginning of the corresponding session

USE OF GENERATIVE ARTIFICIAL INTELLIGENCE	
	You will have the opportunity to use GenAI on many occasions during your studies. However, based on the learning outcomes of this module, in this assessment **you are not permitted** to use GenAI at all. Suspected use of Generative AI will be investigated under the Use of Unfair Means Policy as stated in the Regulations.

NOTE
This piece of assessment should be completed and submitted by the student (or group of students in group work) without assistance from or communication with another person either external or fellow student (outside the group). All sentences or passages cited in the assignment from other people's work should be specifically acknowledged by complete and accurate reference to the author, work and page(s). Failure to abide by the above regulation constitutes use of unfair means (collusion, plagiarism etc.) and will result in a fail mark for this work. It might also invoke disciplinary actions. It is at the instructor's discretion to conduct an oral examination, which will result in the award of the final grade for that particular piece of assessment.

TURN IT IN REQUIREMENT

This piece of assessment is required to be submitted to **turnitin** plagiarism detection software at:

www.turnitin.com

at a date no later than the submission date. This is an absolute requirement for releasing a mark. Brief instructions on how you can set up your profile and submit your work can be found at:
https://help.turnitin.com/feedback-studio/turnitin-website/student/student-category.htm (text)
https://youtu.be/AC3GB-FOMvY (video)

You are going to require:

CLASS ID:	**TBD**	**ENROLLMENT PASSWORD:**	**TBD**

If you have any problems in submitting your work, please contact the course administrator or the module leader.