COMP47350: Data Analytics (Conv)

Dr. Georgiana Ifrim
georgiana.ifrim@ucd.ie

Insight Centre for Data Analytics
School of Computer Science
University College Dublin

2016/17

Lab Work: 2h, 2 parts

- **Problem-Based Learning (1h)**
  - First go at given problem (first week)
  - Second go at given problem + submit report (second week)
- **Practice exercises (1h)**

Lab Work-part1

- **Problem-Based Learning (1h)**

*"*The learning that results from the process of working towards the understanding of a resolution of a problem. The problem is encountered first in the learning process" (Barrows and Tamblyn 1980: 1)

Problem-Based Learning

- Form groups of 5-6
- Decide a recorder for each group
- Work on the given problem and data
- Try to identify what you know and what you need to know to address the given problem
- Summarize discussion (1 page per group)

7 Step PBL Process Guide

Set the Climate

- Assign roles
- Make / review ground rules
- Review thinking and learning processes

1

Read the Problem / Trigger

- Underline and clarify key or unknown terms / facts

2

Define the Kernel of the Problem / Trigger

- Initial ideas

3

Brainstorm

- Ideas / explanations
- Responses and examples in relation to your experience

4

Discuss and Synthesise

- Discuss the problem
- Summarise what you currently know about key themes

5

Formulating Learning Issues

- Name the key issues that you need to study further

6

Independent Study

- Work on learning issues
- Synthesise critically what this means for the problem

7

Professional Practice Debate

- Debate learning issues from the literature and professional practice
- Summarise the learning as it relates to the problem and professional practice

Main Problem

You have been contracted by a motor insurance company to identify fraudulent motor claims using historical data.

Data

Historical data on claimants and claims

	ID	Insurance Ty	Income of Pc	Marital Statu	Num Claimar	Injury Type	Overnight Hc	Claim Amour	Total Claime	Num Claims	Num Soft Tis	% Soft Tissue	Claim Amour	Fraud Flag
2	1	O	0		2	Soft Tissue	No	1625	3250	2	2	1	0	1
3	2	O	0		2	Back	Yes	15028	60112	1	0	0	15028	0
4	3	O	54613	Married	1	Broken Limb	No	-99999	0	0	0	0	572	0
5	4	O	0		3	Serious	Yes	270200	0	0	0	0	270200	0
6	5	O	0		4	Soft Tissue	No	8869	0	0	0	0	0	1
7	6	O	0		1	Broken Limb	Yes	17480	0	0	0	0	17480	0
8	7	O	52567	Single	3	Broken Limb	No	3017	18102	2	1	0.5	0	1
9	8	O	0		2	Back	Yes	7463	0	0	0	0	7463	0
10	9	O	0		1	Soft Tissue	No	2067	0	0		0	2067	0
11	10	O	42300	Married	4	Back	No	2260	0	0	0	0	2260	0

Data Available

A dataset with past claims and claimants (available on Moodle): MotorInsuranceFraudClaimABTFull.csv

Main Problem

- In each lab we discuss a sub-problem of the main problem
- 2 rounds:
  - First go at given (sub)problem (lab1)
  - Second go at given (sub)problem (lab2)

Lab Work – first go at problem

Problem Title	**Does the business problem make sense?**
Problem Description	You have been contracted by a motor insurance company to identify fraudulent motor claims using historical data. You need to decide whether the business problem makes sense as it was originally formulated.
Learning resources	To be identified by students.
Learning outcomes	To be identified by students.

Lab Work – second go at problem

- In the follow-up week, second go at the problem
- Group presentations (max 3mins per group) and debate about relevant questions/answers in relation to posed problem
- By the end of lab, submit final group report (1 page template on Moodle, up to 5 bullet points)

Lab Work: 2h, 2 parts

- **Problem-Based Learning (1h)**
  - First go at given problem (first week)
  - Second go at given problem + submit report (second week)
- **Practice exercises (1h)**

Lab Work-part2

- Practice exercises (1h)
  - Make sure Python3.5 and Jupyer Notebook are installed and running
  - Practice running Python code using Jupyter Notebook
  - Read the Notebook-tips provided on Moodle