

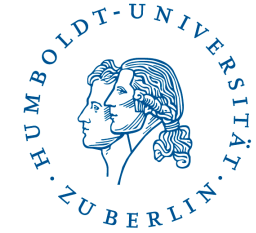
OU LAD

Learning Analytics Dataset

Jakub Kuzilek

<https://www.nature.com/articles/sdata2017171>

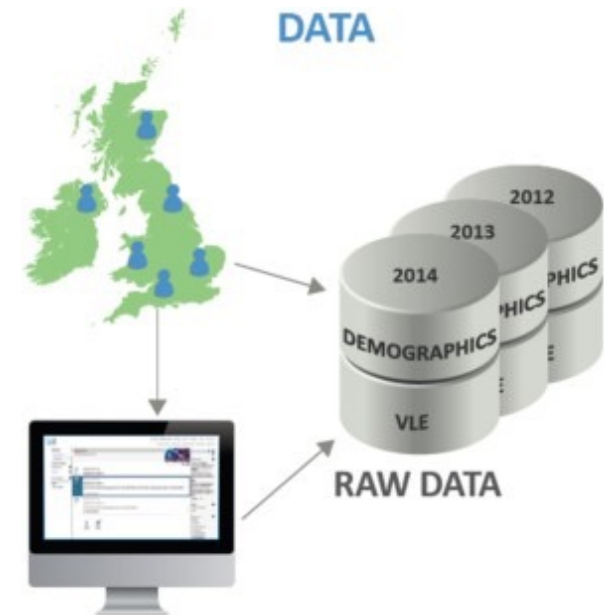
About Open University



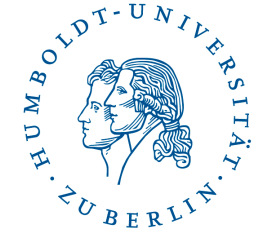
- One of the largest distance based learning universities worldwide
- Approx. 180,000 students
- Teaching delivered remotely via Moodle-like system
- During the degree students participate in several courses (modules)
- 1 module ~ 60 credits
- Wide variety of units/departments exists to support students



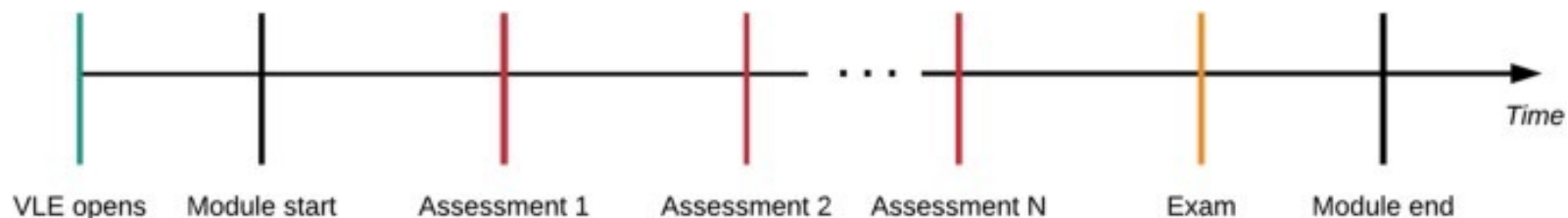
The Open University



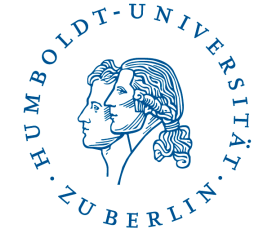
About courses



- Each academic year has two semesters – winter (J) and summer (B)
- Typical semester last approximately $\frac{3}{4}$ of the year
- Students interact with the Virtual Learning Environment, which contains all the learning resources, assessments and tools for contacting their teachers and peers
- Course-presentation covers multiple topics, each usually ends with the assessment => assessments can be considered as milestones in course-presentation

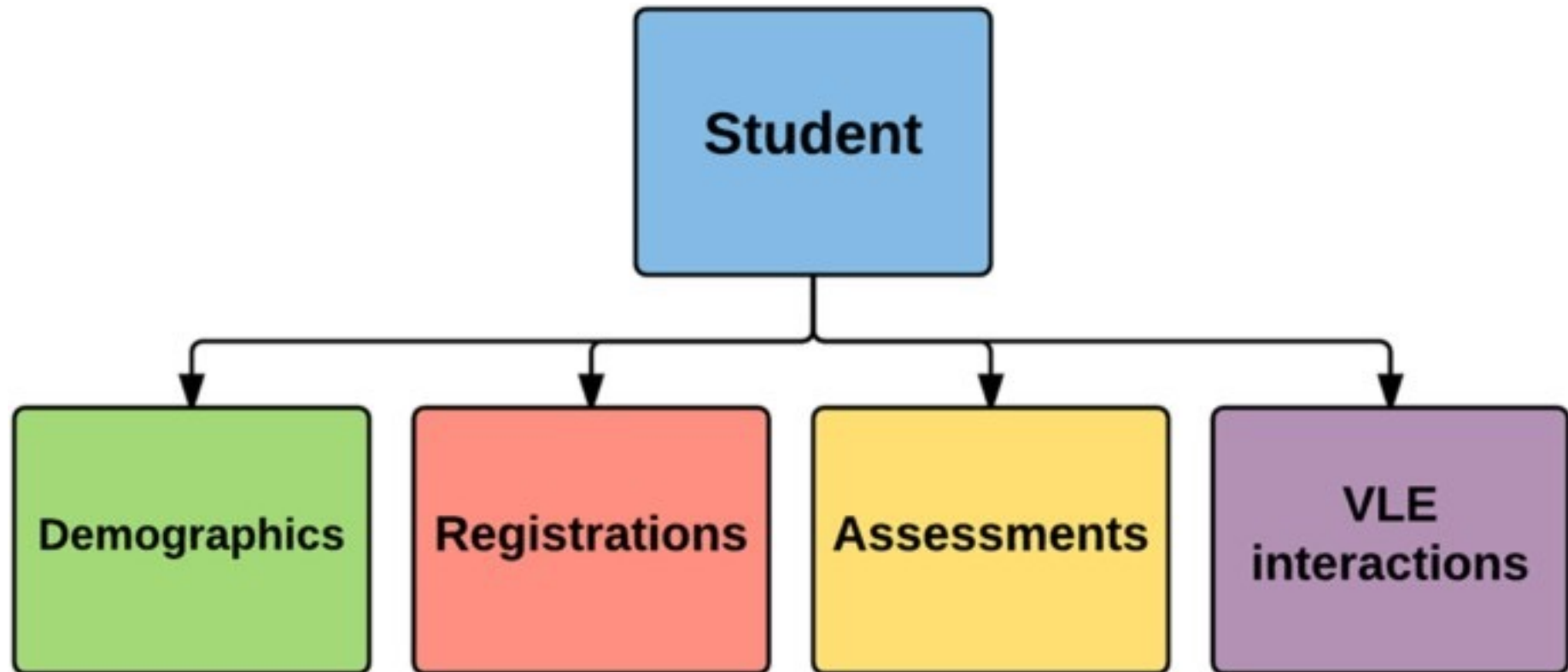
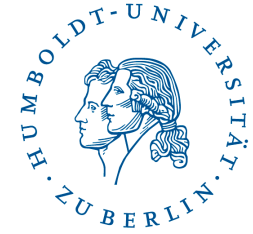


Dataset info

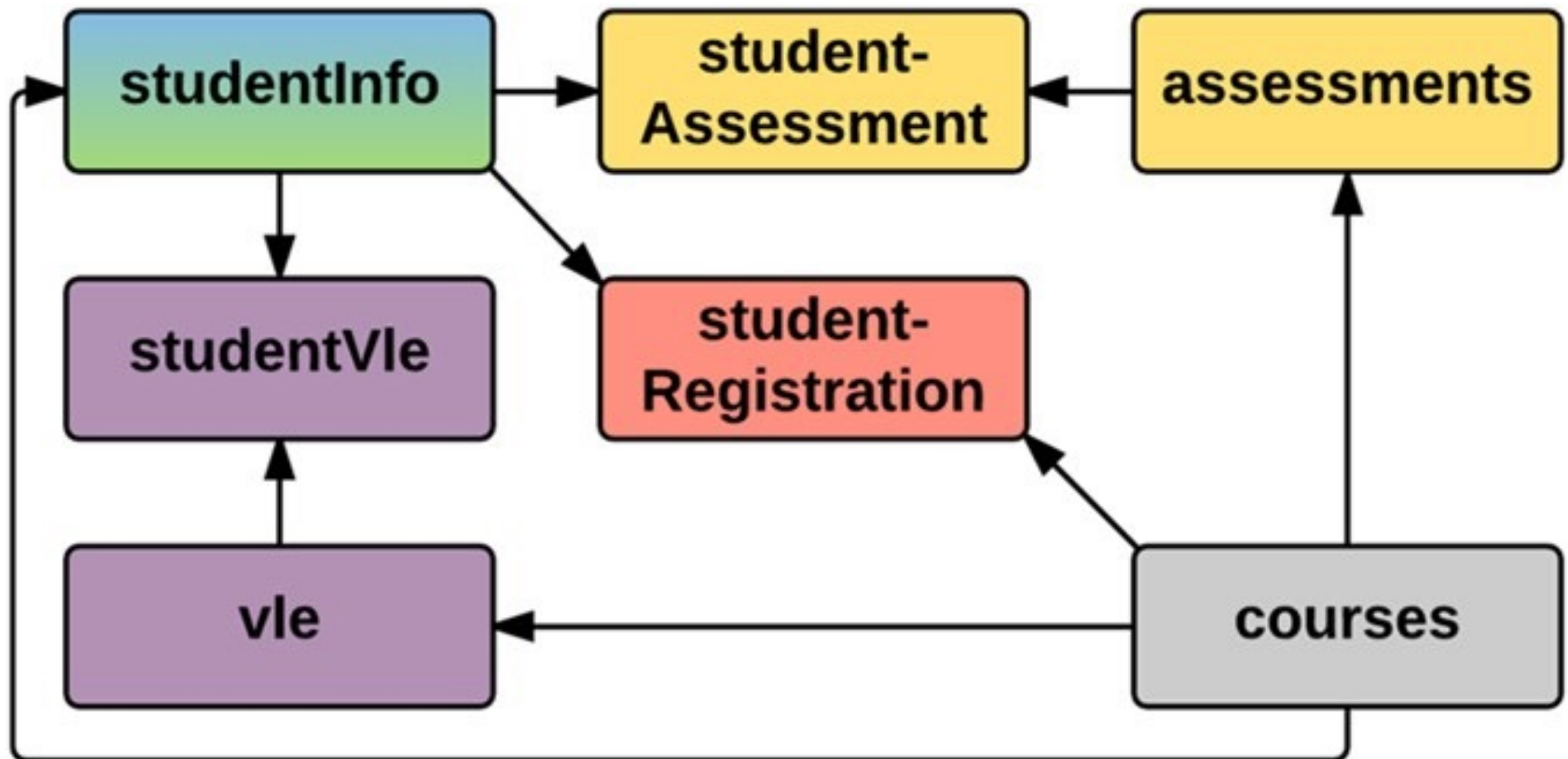
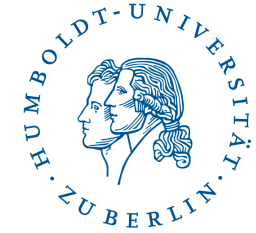


Module	Domain	Presentations	Students
AAA	Social Sciences	2	748
BBB	Social Sciences	4	7,909
CCC	STEM	2	4,434
DDD	STEM	4	6,272
EEE	STEM	3	2,934
FFF	STEM	4	7,762
GGG	Social Sciences	3	2,534

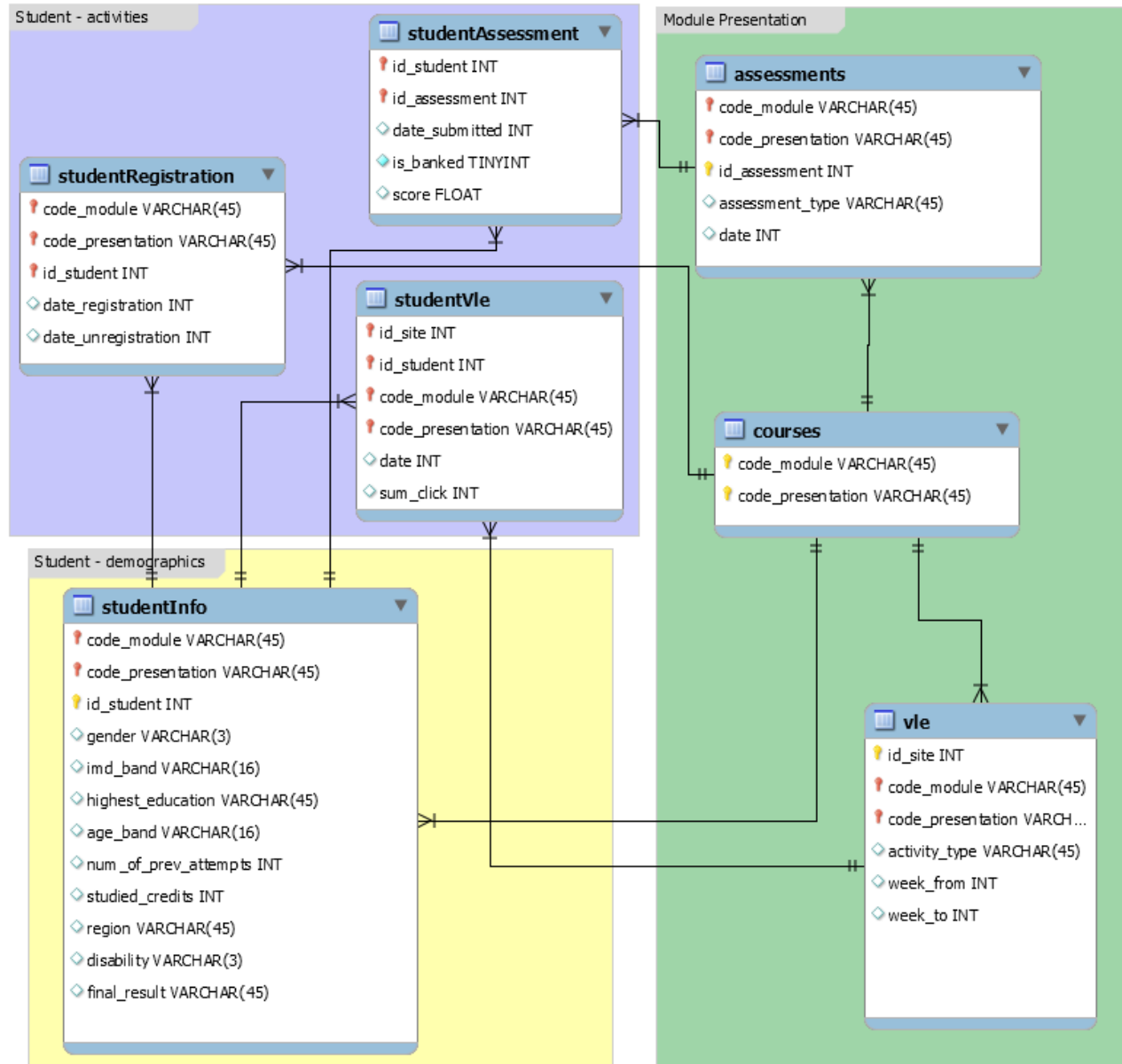
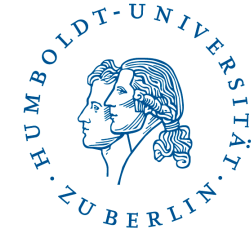
OULAD structure



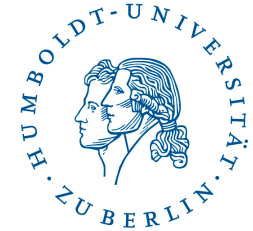
OULAD structure



E-R diagram

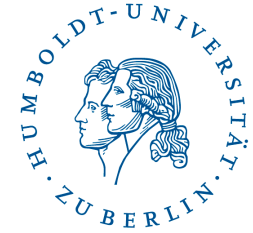


student_info



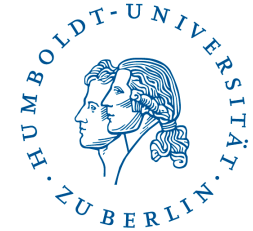
- ***code_module*** - module identification code on which the student is registered.
- ***code_presentation*** - presentation identification code during which the student is registered on the module.
- ***id_student*** - the unique student identification number.
- *gender* - student's gender.
- *region* - the geographic region, where the student lived while taking the module-presentation.
- *highest_education* - the highest student education level on entry to the module presentation.
- *imd_band* - the IMD band of the place where the student lived during the module-presentation.
- *age_band* - a band of student's age.
- *num_of_prev_attempts* - the number of how many times the student has attempted this module.
- *studied_credits* - the total number of credits for the modules the student is currently studying.
- *disability* - indicates whether the student has declared a disability.
- *final_result* - student's final result in the module-presentation.

assessments



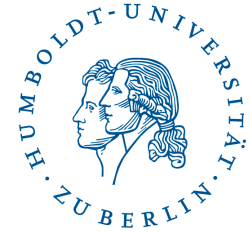
- ***code_module*** - module identification code, to which the assessment belongs.
- ***code_presentation*** - presentation identification code, to which the assessment belongs.
- ***id_assessment*** - assessment identification number.
- **assessment_type** - a type of assessment. Three types of assessments exist—Tutor Marked Assessment (TMA), Computer Marked Assessment (CMA) and Final Exam (Exam).
- **date** - information about the cut-off day of the assessment.
- **weight** - the weight of the assessment. Typically, Exams are treated separately and have the weight equal to 100%; the sum of all other assessments is also 100%.

student_assessment



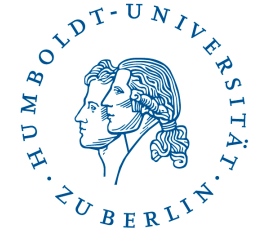
- ***id_assessment*** - the assessment identification number.
- ***id_student*** - the unique student identification number.
- *date_submitted* - the day of assessment submission.
- *is_banked* - the status flag indicating that the assessment result has been transferred from a previous presentation.
- *score* - the student's score in this assessment. The range is from 0 to 100. The score lower than 40 is interpreted as Fail. The marks are in the range from 0 to 100.

courses



- ***code_module*** - code name of the module, which serves as the identifier.
- ***code_presentation*** - code name of the presentation.
- *length* - the length of the module-presentation in days from module start date to module end date.

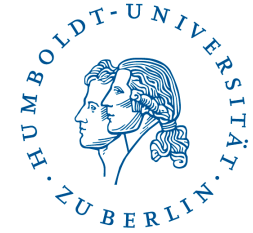
student_registration



- ***code_module*** - the module identification code.
- ***code_presentation*** - the presentation identification code.
- ***id_student*** - the unique student identification number.
- ***date_registration*** - the day of student's registration for the module presentation.
- ***date_unregistration*** - the day of student unregistration from the module presentation. Students, who completed the course have this field empty. Students who unregistered have Withdrawal as the value of the final_result in the studentInfo table.

- ***id_site*** - the identification number of the material.
- ***code_module*** - the identification code for the module.
- ***code_presentation*** - the identification code of the presentation.
- *activity_type* - the role associated with the module material.
- *week_from* - the week from which the material is planned to be used.
- *week_to* - the week until which the material is planned to be used.

student_vle



- ***code_module*** - the module identification code.
- ***code_presentation*** - the presentation identification code.
- ***id_student*** - the unique student identification number.
- ***id_site*** - the VLE material identification number.
- *date* - the day of student's interaction with the material.
- *sum_click* - the number of times the student interacted with the material.