**Advanced Databases**

**Assessment and Module Handbook**

2020/21 Students

**Level 6, Block A**

(20 Credits)

Student Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Email Address \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Course \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Group \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Module tutor \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***Communication Protocol:*** *module staff will reply to student questions within a reasonable time but this will normally be within office hours only. Students are advised to check this Handbook and also to see if there are any online announcements or FAQ answers that deal with their enquiry before contacting staff.*

Contents

[1 What this Module is about 3](#_Toc46131373)

[1.1 Welcome to Advanced Databases! 3](#_Toc46131374)

[1.2 Module Learning Outcomes : 3](#_Toc46131375)

[1.3 Module Learning Activities 3](#_Toc46131376)

[1.4 Graduate Attributes Developed and Assessed 4](#_Toc46131377)

[2 Schedule of Work 5](#_Toc46131378)

[3 Key Resources to Support Learning (see reading list on VLE) 9](#_Toc46131379)

[4 AD - ASSIGNMENT REQUIREMENTS 12](#_Toc46131380)

[4.1 Introduction to the Assignment using the FlyU case study 12](#_Toc46131381)

[4.2 Assignment 1: Data Mart (DM) – Design and ETL considerations [50%] 13](#_Toc46131382)

[*4.3* Assignment 2: Data Mart (DM) data staging, ETL, implementations and considerations [50%] *Formative feedback for task 3 is available in class in week 11* **Error! Bookmark not defined.**](#_Toc46131383)

[Appendix A 15](#_Toc46131384)

[Appendix B. 16](#_Toc46131385)

[ASSIGNMENT 1 MARKING SCHEME – worth 50%. mark out of 100% 18](#_Toc46131386)

[ASSIGNMENT 2 MARKING SCHEME – worth 50%. mark out of 100% 20](#_Toc46131387)

[4.4 Reassessment 24](#_Toc46131388)

[4.5 Feedback 24](#_Toc46131389)

[5 Understanding Your Assessment Responsibilities 25](#_Toc46131390)

# What this Module is about

## Welcome to Advanced Databases!

On behalf of the module team, I would like to welcome you to this module, which we hope you will find both challenging and rewarding.

The module is based around the use of data for business decision making and business intelligence solutions and data analytics dashboards and applications. We will follow a database mart/ warehouse development lifecycle by firstly looking at what we want to investigate in the data mart along with the types of reports (requirements). With this information we can identify the data required and use it to design an appropriate data model (design), part of the design will involve checking (and dealing with) the data quality. The implementation will involve the creation of a data mart data model and writing SQL based ETL scripts to ‘extract, transform and load’ the data from the original data sources into the data warehouse model. At this point we can produce some reports and consider the best visualisation tools for this.

Along the way topics such as data integration, data quality, the data dictionary, data maintenance, and data ethics will be discussed. The module has a good mixture of database expertise and business understanding – you will write SQL and you will discuss business strategies! As your module tutors, we aim to provide you with a coherent set of learning opportunities, which will enable you to develop your skills and knowledge in databases.

We hope this will be a valuable learning experience for you.

Jackie

Module Aims

* This module aims to build understanding and practical capabilities of database technologies for effective management and utilisation of an organisations data resource.
* The principles, techniques and concepts of data integration, extraction and data quality when implementing a database system are addressing theoretically and practically.
* Your knowledge of technical and practical development and the end-purpose of using a database system for decision support making will be developed.

## Module Learning Outcomes :

On completion of this module the student should be able to:

1. Demonstrate a critical understanding of the theories underpinning a range of database design and management issues, including methods, technologies and emerging trends
2. Critically evaluate the role of development tools in the design, development and management of database systems
3. Assess the organisational context, roles and tools for effective utilisation of an enterprise’s data resource.

## Module Learning Activities

Keynote lectures will be used to develop knowledge and understanding and to provide a discussion of students’ own research finding.

Tutorial sessions will be lab-based and will be mainly problem or enquiry based, allowing students to analyse, evaluate and discuss these technologies using case study scenarios, and to use and appraise applications and technologies. These activities will be both individual and group-based.

Students will discuss and present ideas to their peers, enabling tutor and peer feedback.

Independent research and application will be expected as students will need to read around the subject in order to gain a wider understanding of the theory application of the technologies covered.

## Graduate Attributes Developed and Assessed

* The Enterprise graduate attribute is both developed and assessed via problem based activity.
* The digital literacy graduate attribute is both assessed and developed via the investigation and selection of appropriate tools for the activities.
* The global outlook graduate attribute is developed in that students are required to consider the case study activities for an international market.

# Schedule of Work

You will need to revise the material from L4 and L5 and remind yourself of database design and modelling. An introduction to database systems by CJ Date Chapters 1-4 is good – although others listed in the reading list are preferred by some students. There is also material on the VLE under Unit 1.

General background and appreciation to this module can be found in the book Big Data: A Revolution That Will Transform How We Live, Work and Think[Viktor Mayer-Schonberger](http://www.amazon.co.uk/s/ref=ntt_athr_dp_sr_1?_encoding=UTF8&field-author=Viktor%20Mayer-Schonberger&search-alias=books-uk&sort=relevancerank)  and [Kenneth Cukier](http://www.amazon.co.uk/Kenneth-Cukier/e/B00C47ZFSY/ref=ntt_athr_dp_pel_2) . Copies are in the library.

|  |  |  |
| --- | --- | --- |
| Unit 1: Decision Support systems, Analytics and the Data Warehouse.  (overview) | | |
| 1 | **Prework – go through the L4 Fundamentals of db workbook**  **Prework: Go through the SQL practice from L5**  **Prework - Reading: Data Warehouse - Data Model Overview chapter in Turban p. 1-6**  **Remote Lecture: Overview of the Module (video)**  **Introduction to Decision Support and BI and the role of a Data warehouse in this (video)**  Face to face and remote tutorial:   * Introduction to our learning environment * Overview of the module * The role of Decision Support, BI systems, the data warehouse, BI solutions   **Practicals ERD recap:**  **Exercise on ERD and SQL Revision based on for assignment case study**  **Meet a new ER model – the star schema (SS)**  **SQL on the SS model** | Evidence of self study and practice, also formative work for the assignment.  Upload: Completed class SQL exercises based on the ERD and the star schema (SS) model  By: 12th October 2020  This contributes to the final grade |
| Unit 2 – OLAP and OLTP – Decision support  (Understand the Business and Data) | | |
| 2 | **Prework - Reading: Data Model Overview chapter in Turban p.7-14**  **Prework – Consider investigations for**FlyU**and the reports to support this.**  Remote Lecture: Data and Design for Decision support Operational systems:   * DW requirements/OLAP and OLTP (video) * The star schema (video) * DW Approaches - architecture – Inmon’s Information Factory, Kimball’s Dimensional DW, stand alone Data Marts, Data lake, enterprise DW, virtual DW (video) * Methodologies – CRISP, Data Life Cycle, TSDP, PPDAC (video)   Face to face and remote tutorial:  Review of uploads from Unit1  Exercise on OLAP, OLTP and types of reports.  **Practical: Identify an investigation for FlyU, stakeholders, users, reports, create a table to show where the source data for these reports is held. (We return to this in Unit 4)** | Upload: Practical task – getting started with assignment.  By: 19th October 2020  This contributes to the final grade |
| Unit 3 – Data Marts and Star schema  (Understand the Business and Data) | | |
| 3 | **Prework - Reading: Data Warehouse - Data Model Overview - finish chapter**  **Prework – complete all exercises from Unit 2 on OLAP, OLTP and getting started on assignment**  **Remote Lecture – Data Marts and star schemas (video)**  **Slowly Changing Dimensions (video)**  Face to face and remote tutorial:  Practical: Designing star schemas and understanding the data (integration and sources)  Take it further – data maintenance and Slowly Changing dimensions | Upload a document including:   * your 4-5 reports, * the star schema to support these and * the code of the forward engineered star schema using QSEE.   By: 26th October 2020  This contributes to the final grade |
| Unit 4: ELT - Data Integration, Data Quality, Data Prep | | |
| 4 | **Prework: Return to the practical exercise in Unit 2. Rework this.**  **Add to this:**  **A plan of the data that will be in the reports – actually work it out from the raw data sources. You could also play around with SQL to help with this (save the SQL).**  **Remote Lecture: ETL- Data integration, data quality and data transformation (videos)**  Face to face and remote tutorial:  Practical: Demo and practice of merging (integrating) the data required for the CLAIMS data mart.  Demo and Practice of dealing with data quality issues.  Documenting the ETL  Take it further – use of log tables, triggers, sequences, PL/SQL | Start on integrating data and dealing with DQ issues for your assignment.  No upload. |
| Unit 5 – Data Transformation and ETL (ETL)  (Data Prep and data model) | | |
| 5 | **Prework: Work on the assignment, as by now you should be asking questions about PKs/FKs, fact id’s, number of dimensions …**  **Remote Lecture – ETL/SS - questions you may be asking, things to consider**  Face to face and remote tutorial:  Practical: Demo and practice data transformations for the CLAIMS data mart.  Demo and Practice of loading data into dimensions and fact table.  Take it further – look into SCD | Upload:  The scripts so far:  Script 1 – creates the SS tables (done with QSEE)  Script 2 –  Merges data required into tmp tables.  Populates the dimension tables (time\_dim and 2 more)  Populates the fact table (one measure only)  By: 9th November 2020  This contributes to the final grade |
| These next Units are exploring the BI/user side of the data warehousing project. Each Unit presents a different aspect of the project. After undertaking the practical work – we critique, reflect, evaluate the different DW approaches (Information Factory, Bus Architecture, Stand Alone Data Mart) | | |
| Unit 6: Using tools, tableau for data prep and data analysis  (data prep, data analysis) | | |
| 6 | **Remote Lecture :**  **Tableau for prepping data (video)**  **The assignment (video)**  Face to face and remote tutorial:  How to evaluate a technical solution (video) as well  Using tableau for data prep (video)  Using tableau for data analysis  Review of the data prep phase of the Data project lifecycle  Evaluation of the DW approaches (for prepping and storing data to use for BI)  Catch up and one to one conversations regarding assignment tasks so far. | No upload |
| Unit 7: Extracting data from the Data Mart for analysis – in Access/Excel and using Oracle Apex  (data analysis) | | |
| 7 | **Remote Lecture – OLAP (video)**  Face to face and remote tutorial:  OLAP, extracting data from a ‘data warehouse’ and using Excel as an analysis tool for OLAP using pivot tables.  The detail table  Review of the data extract and communication of the findings stage of the Data project lifecycle.  Evaluation of the DW approaches (to support OLAP)  Take it further, role of documentation, good visuals, story telling and narratives | Make research notes on dashboards, what is their role why are they used? Who are the users? Are dashboards OLAP? Where does their data come from?  No upload- please prepare for class |
| Unit 8: Extracting data from the Data Mart for analysis – for dashboards  Data Maintenance | | |
| 8 | **Prework – Prepare 2-3 SQL reports based on your star schema model**  **Remote Lecture :**  **The role of dashboards (video)**  **Data Maintenance (video)**  Face to face and remote tutorial:  Dashboards  The sample reporting application and building a dashboard in apex  Data Maintenance consideration – updating the DW and enterprise reports  Evaluation of the DW approaches (data life cycle and project approaches)  Extra - **Business Performance Management (video)** | No upload  Work on assignment tasks |
| **AD assignment hand-in. Upload to the VLE by Friday 11th December23:00** | | |

# Key Resources to Support Learning (see reading list on VLE)

Your general understanding and application of a subject area will improve the more you ‘read around’. In the final year of studies, we expect to be confident in researching literature on topics and appreciating the varied views and insights this may offer – so the suggestions below are just that, suggestions and recommendations.

Data and databases – of interest

* Jeffery Deaver, **The Broken Window** published 2008 by Hodder & Stoughton ISBN 978 0 340 99370 5.This book is a thriller, the criminal uses information and data mining to frame his victims. It has some interesting observations on the importance of data, the volumes of data, meta-data, data mining and forensic evidence. Available in the library.
* TV programs such as **CSI, Numb3rs** often offer some good insights into how data/statistics are used to identify suspects or ‘profile’, gain data mining. You need to pick your episode though.
* **Youtube** shouldn’t be overlooked for videos of lectures or ‘how to’s (eg normalization, data mining)
* The **Guardians ‘data’ blog** has some interesting data analysis stories. <http://www.guardian.co.uk/data>, their technology sections also includes some good podcasts.
* **Computer Weekly** is available in the library and on: <http://www.computerweekly.com/Home/>. It is The Computing paper, lots of current information, white papers and jobs. A good place to research a company (with respect to their IT systems) before you go for an interview.
* **Oracle Technet Web Site** (FIES 1999 Comment: OTN - the ORACLE TECHNOLOGY NETWORK - contains a maze of useful material, loads of useful ORACLE stuff. Anyone can register.) <http://technet.oracle.com/>

Note that when referencing sources such as those above, some may be subjective (opinion rather that fact).

**More specific material for the module, I have detailed material under the areas covered on the module. They are all available in the library. Many of the books will have earlier editions in the library – these are usually fine. Please look yourself for books you like and let me know any that you find useful.**

1. **ERD,LD,SQL**

**Database generics:** You will have already covered these areas. I have put some revision material on the VLE (lecture notes, exercises). As before there are some excellent books that are worth looking at to help clarify and expand your understanding. For example:

Hoffer, J. A. & Prescott, M. B. & Topi, H. (2016) **Modern Database Management,** 12th ed. London, Pearson Education.

Elmasri, R. & Navathe, S. (2015) **Fundamentals of Database Systems, 7**th ed. London, Pearson Addison-Wesley.

**Connolly, T. M. & Begg, C. E. (2014)** Database Systems – *A practical approach to Design, Implementation and Management*, **6th ed. Harlow, Addison Wesley.**

**Shah, N. (2004)** Database Systems Using Oracle – A Simplified Guide to SQL and PL/SQL**. Prentice Hall.**

**Date, C. J. (2004)** An Introduction to Database Systems**, 8th ed. London, Addison- Wesley.**

**PL/SQL :** You should aim to develop your SQL into PL /SQL by now and start to use functions, packages and procedures.

**Steven Feuestein’s Site** (is a good site for PL/SQL and code generation.) <http://www.stevenfeuerstein.com/>

Oracle PL/SQL Programming Paperback  [Steven Feuerstein](https://www.amazon.co.uk/Steven-Feuerstein/e/B001IGNQSS/ref=dp_byline_cont_book_1)

Casteel J, **Oracle 9i Developer: PL/SQL programming**, Thompson

**Apex development:** We don’t explicitly do application development in Apex on this module. Although you may choose to do some to illustrate your findings in the assignment. You may also need a refresh for the final year project.

Gault et al , (2015) **Beginning Oracle Application Express 5**,  3rd ed. 2015 edition

John Edward Scott Spendolini,  **Oracle Application Express**

1. Data Dictionary, Data marts and warehouse, Data Mining, Case tools

Also look at: [www.dba-oracle.com/concepts/data\_dictionary.htm](http://www.dba-oracle.com/concepts/data_dictionary.htm) for an overview of data dictionary concepts.

On the VLE are some papers I thought were relevant (you will find many papers /journals and white papers on data cleansing, data marts and warehouses, data mining and case tools).

Books on data warehousing:

Adamson, Star Schema: The complete reference

These two are my favourite they cover the life cycle, design, ETL and applications ( I have a personal preference for the first):

[Efraim Turban](https://www.amazon.co.uk/Efraim-Turban/e/B001ILKERO/ref=dp_byline_cont_book_1), [Ramesh Sharda](https://www.amazon.co.uk/s/ref=dp_byline_sr_book_2?ie=UTF8&text=Ramesh+Sharda&search-alias=books-uk&field-author=Ramesh+Sharda&sort=relevancerank), Dursun Delen (2010). Decision Support and Business Intelligence Systems

[Efraim Turban](https://www.amazon.co.uk/Efraim-Turban/e/B001ILKERO/ref=dp_byline_cont_book_1), [Ramesh Sharda](https://www.amazon.co.uk/s/ref=dp_byline_sr_book_2?ie=UTF8&text=Ramesh+Sharda&search-alias=books-uk&field-author=Ramesh+Sharda&sort=relevancerank), Dursun Delen (2014). Business Intelligence and Analytics: Systems for Decision Support.

The following are more specific to data models (including star schemas and data warehousing).

Silverston, L. The Data Model Resource Book. New York, Wiley.

Kimball, R. (2008) The Data Warehouse Lifecycle Toolkit. New York, Wiley.

Bert Scalzo, Oracle DBA Guide to Data Warehousing and Star Schemas, Prentice Hall; 1 edition (June 14, 2003)

George M. Marakas Modern Data Warehousing, Mining and Visualisation, Prentice Hall

Hobbs, L. Hillson, S. Lawande S. Smith P. Oracle Database 10g Data Warehousing. Elsevier\*

\*Good for first assignment question

3) Data Maintenance, Data mining, DBA

Again many of the books above (in general databases and data mart/warehouse sections) will include information on data mining and DBA concepts.

Data maintenance is a more generic term that covers the updating/deleting and inserting of data. The exercise we do is an Extract and Load of data.

Reading around on ETL software and scripting and data cleansing will aid your understanding of this area.

4) Data Analysis and visualisation

Get started with OLAP using Excel or PowerBI.

<https://powerbi.microsoft.com/en-gb/documentation/powerbi-sample-downloads/#download-the-power-bi-sample-workbooks>

Learn Excel Essentials and Expert, the smart way.

<http://excelcentral.com/> (book(s) available in the library)

[David McCandless](https://www.amazon.co.uk/David-McCandless/e/B002HYXAUY/ref=dp_byline_cont_book_1)  (2012) Information is Beautiful.

Stephen J. Dubner and Steven D. Levitt, Freakeconomics

**We use Oracle Apex**

Apex has significant help, tutorials and documentation available via the ‘help’.

# Design and considerations of a Data Mart - ASSIGNMENT REQUIREMENTS

## Introduction to the Assignment using the FlyU case study

FlyU is an airline company. They keep records of the flights that leave each airport. They record planned departure and arrival times, actual departure and arrival times, from and to each airport. Information on which passengers are on each flight is recorded, some passengers will need a number of flights to reach their destination – this is known as a trip and also includes details of return flights. Also recorded is the lead and deputy pilot for the flight. The tail number is the id for the plane (similar to a car registration number).

Complaints raised by passengers are recorded against flights, passengers are required to raise a complaint against each flight if they miss connections due to delays.

The system provides information for the airline on flight delays which they use in their planning and customer complaints service.

You have been given a number of data sources.

1. **The FlyU complaints system (oracle database)**

A database management system used by the FlyU office to record details of complaints raised (see Appendix A).

and

1. **Spreadsheets with information useful to the airline (see Appendix B)**

* Weather Information
* Detailed flight information for 2017 and 2018

Your role is as an analyst/ developer on a Data Mart (DM) project to support the design, analysis and collection of information relating to FlyU.

## Data Mart (DM) – Design and ETL considerations

**FlyU have a number of KPI’s (Key Performance Indicators) they would like to consider. These include:**

**KPI 1: Deliver a quality service**

This KPI is concerned with the quality of their service to provide a good flight experience to customers. In order to investigate the quality of their service they wish to investigate the outcomes of flights. Examples of the types of reports they would like are:

* Number of flights which are delayed per year
* Number of cancelled flights per year

**KPI 2: Ensure customer satisfaction**

This KPI is concerned with ensuring that customer complaints are dealt with adequately and do not cause customers to not use them again. Examples of the types of reports they would like are:

* Number of customer complaints per year
* The amount of compensation per year

**KPI 3: Grow the company**

This KPI is concerned with developing the business, buying new aeroplanes, flying from new airports and flying more. Examples of the types of reports they would like are:

* Number of flights per airport per year
* Number of airports per year
* Number of airmiles per year

Choose one of the KPIs above to focus your assignment around. Address all the tasks with respect to this KPI.

Use the tasks below as a template for the report you hand in.

***Formative feedback for task 2 is available in class sessions and via email***

**Task 1: Data Mart (DM) star schema design for your chosen KPI**

* Identify 3-5 reports\* that your star schema will support.
* Document the star schema (SS) design model to support these reports – use QSEE
* Use the data dictionary template from tutorials to document the data model the project.
* Select one of the reports\* you have suggested. Illustrate the expected data in the star schema to support the report - use Excel (or oracle or similar) to do this and add a few rows.

**[20 marks]**

**Task 2: Pre-processing: Gather the data sources prior to ETL**

1. **Collect ONLY the data you require to populate the star schema (source data for DM) into the data staging area. [5 marks]**

Your Apex oracle account is effectively the ‘data-staging’ area for the exercise.

To collect the data you need for your Data Mart you will need to run the customer complaints script and upload the data from the spreadsheets. You may choose the amend the spreadsheets before upload to apex or select from them and create new spreadsheets as an initial staging area.

1. **Star Schema set up (DM environment) [5 marks]**

Use QSEE to forward engineer the database for the star schema (SS) database you have designed. Create and run a script to create the data mart tables (edited as appropriate).

Include QSEE generated script(s) as part of your upload along with screen shots as evidence of the code running successfully and documentation of any changes you have made or problems you encountered.

*See marking scheme for more direction.*

**Task 3: Extract, Transform and Load (ETL) script to populate the Star Schema (DM) with data**

The ETL is a script that puts the data into the DM tables. It does this by extracting the data from the original sources, transforming the data as required and then loading them into the DM tables.

Write an ETL script to:

* populate one - two dimension tables,
* the time\_dimension table and
* 10-20 rows of the fact table with measure(s).

To do this, identify one of the reports to support (ideally the one you have already planned the expected data for task 1). Your script should deal with 2-3 data quality issues, 1 transformation and include at least 1 measure/calculation for the FACT table. **[20 marks]**

Perform and provide evidence that you have successfully completed these tasks (via screen shots which show your student id or evidence of successfully run scripts, this work should be done in your own University apex account). Documenting any changes you have made or issues you have encountered.

## Data Mart (DM) – OLAP, dashboard DW approach

**Task 4: Data Analysis/OLAP/Mining Investigation**

1. Upload one of the case study spreadsheets into MS Excel (or tableau). Create a pivot table and produce some interesting (and appropriate) reports using the charts and visualisations functionality. *Use literature of data analysis, business intelligence and OLAP to support and drive this task*. You may include (or discuss) external data as well. **[15 marks]**

OR

1. Using Apex create a dashboard for FlyU. *Use literature on data analytics and dashboards to inform the design*. The dashboard should support your chosen KPI for Assignment 1. You can use the data from the DM tables or the source data tables. **[15 marks]**

**Task 5:** Consider and reflect on data warehousing approaches with respect to FlyU. Draw on the models of Kimball and Inmon in considering solutions for supporting an enterprise (business) and the business processes and functions. (800 words)

**[15 marks]**

S*ee marking scheme for more direction.*

**Evidence of portfolio uploads [20 marks]**

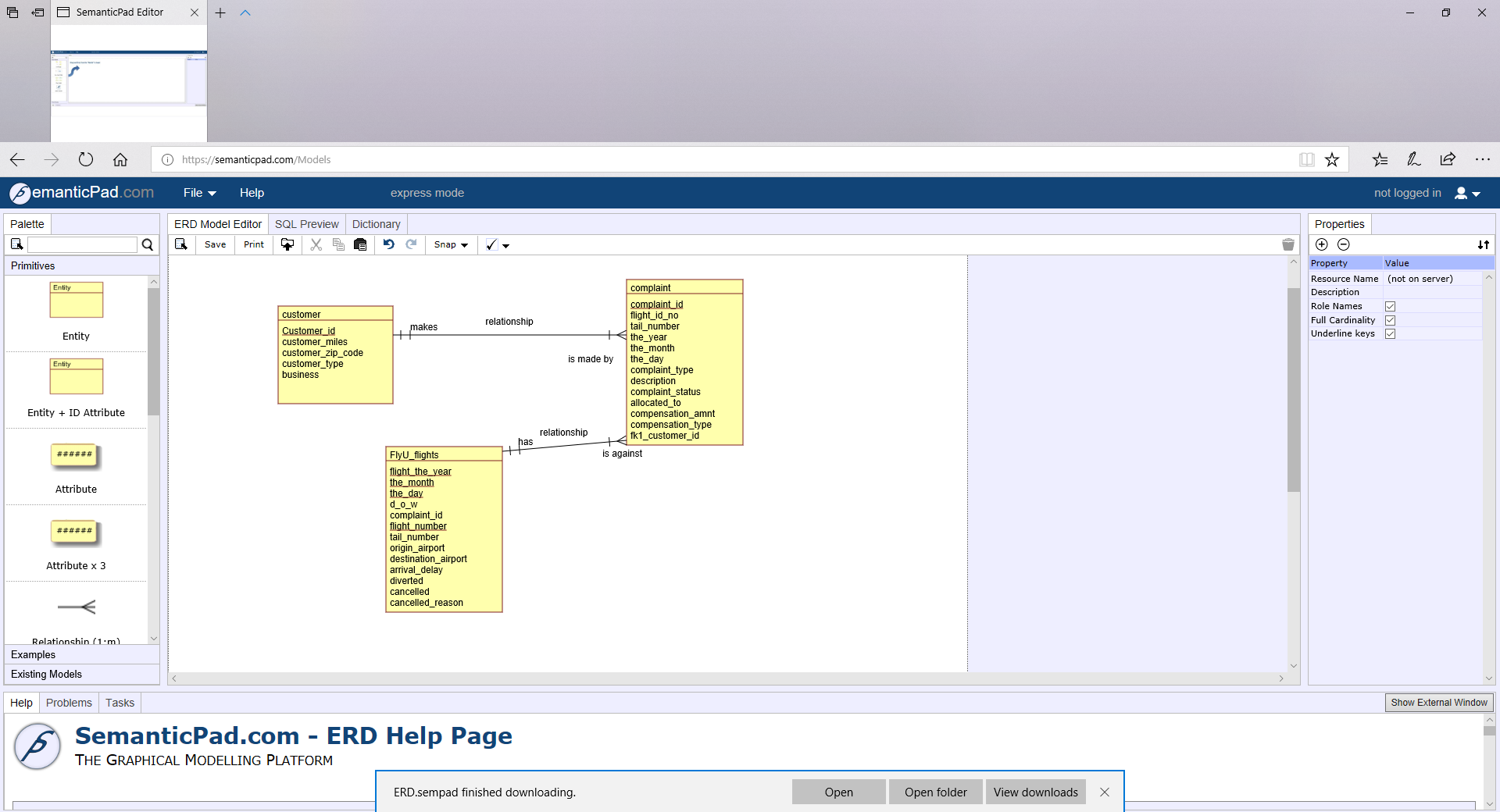
**Assignment upload: Please upload a word report addressing tasks 1-5, include any code as an appendix to this document.**

**Upload to VLE by Friday 11th December 2020 23:00.**

### Appendix A

FlyU

The system below holds the information on all complaints made to the airline.

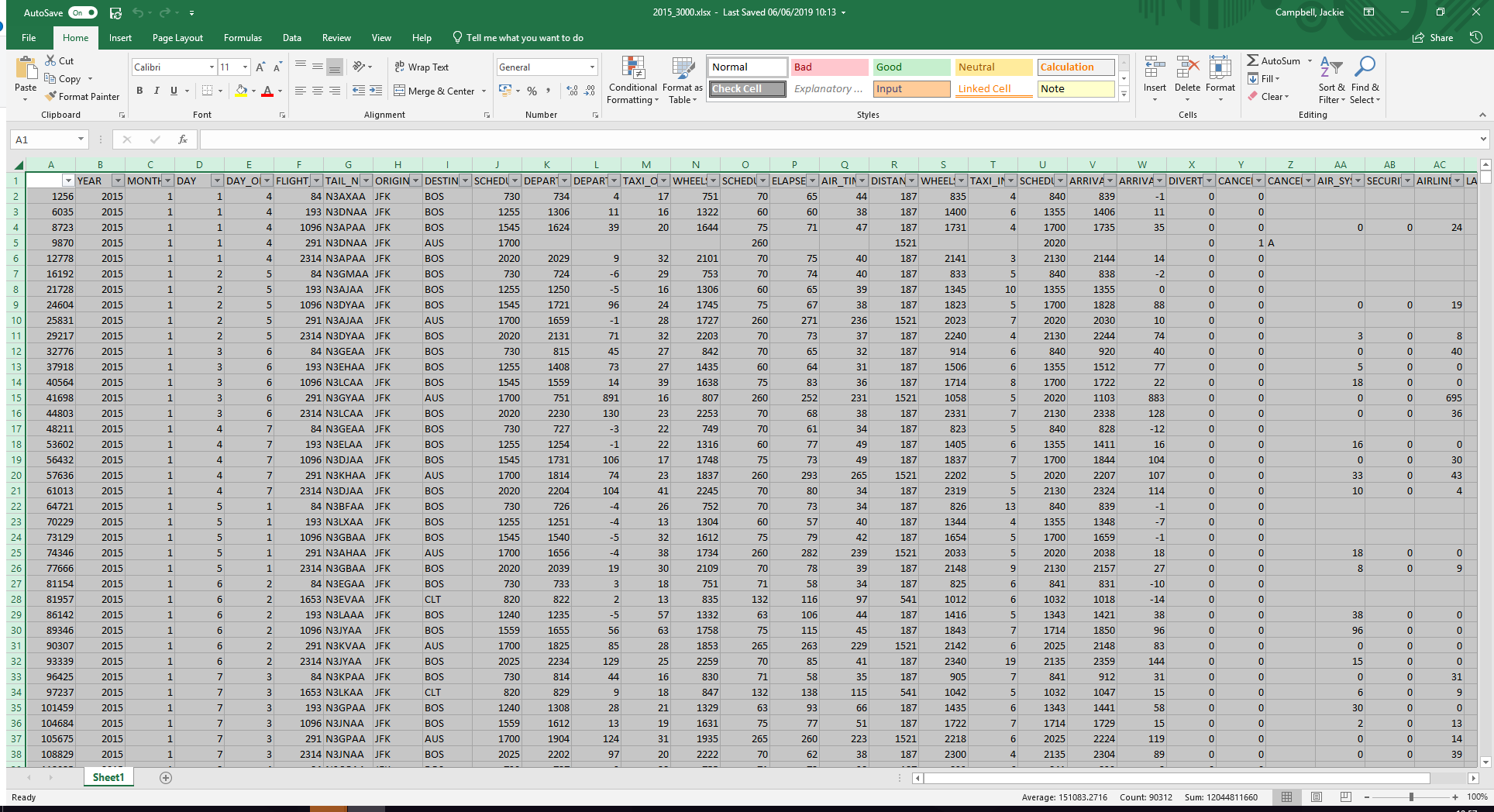


The above was created in semanticpad is the link: <https://www.semanticpad.com/>

Use "express mode" for now. This is the new version of QSEE.

### Appendix B.

FlyU – Flights spreadsheet example



### 

### ASSIGNMENT MARKING SCHEME mark out of 100%

*Student no: Group: Date:*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ***Criteria*** | ***>85%*** | ***Level: 70%+ 1st*** | ***60 – 69% 2:1*** | ***50 – 59% 2:2*** | ***40 – 49% 3rd*** | ***<39% Fail*** | ***<30%*** |
| **Task 1:**  **DM Star schema design**  **20 marks** | Exceptional work as for first and also showing unique/ self-taught /outstanding application of tools, research, understanding | 3-5 reports of DM type listed and complement the chosen KPI.  Star schema fully specified and meets requirements specified design decisions discussed and justified via the data dictionary(s).  All attributes correctly documented.  Supports functionality and reports as documented.  Design considerations (Granularity design decisions and reflected in star schema design).  The data is illustrated for each dimension and the fact. With a least a few rows in each, M:N.The data is correct based on given case study data.  Shows full understanding.  Clear evidence of application of literature to the assignment tasks | Star schema fully specified and meets requirements, design decisions discussed and justified.  Mostly as for a first, some areas ambiguously defined, missing or generally simple, however excellent understanding evident.  Literature included ‘matched’ rather than applied. | Star schema mostly specified and meets most requirements, some design decisions discussed and justified via a data dictionary.  The reports and examples are very similar to those given in class or in assignment specification.  Generally well done with no major flaws.  Literature included, described rather than used to inform the tasks. | Star schema mostly specified and meets most requirements via a DD.  There are major flaws in the design, considerations or data illustrated.  Some literature. | Star schema doesn’t support queries. Little understanding demonstrated.  Very little data, or incorrect data.  Little understanding demonstrated. | Work evidences very little understanding in the field of study. Very little done or not relevant to task |
| **TASK 2: DM set up (SS and OLTP) (5 marks)**  **SS- FACT table and 2 other tables as a minimum (5 marks)** | Excellent database produced. Evidence that CASE tool has been used. All aspects of the CASE database design issues addressed & dealt with.  *Data from both databases uploaded to Apex. Thorough and considered, only data required to support reports have been uploaded. There may be some pre-processing in excel. This is documented.* | Good database. Evidence that CASE tool has been used. Most aspects of the CASE database design issues addressed & dealt with.  *Data from both databases uploaded to Apex. Thorough and considered, checked using appropriate techniques. Understanding and documentation of only including data required.* | Database produced. Using CASE tool .Some aspects of the database design addressed & dealt with.  *Data from both databases uploaded to Apex. Little discussion or explanation.* | Db produced. Little evidence that CASE tool have been used. Few aspects of the database design addressed & dealt with.  *Data from both databases uploaded*.  *Little or no discussion or explanation.* | Little database, or evidence of CASE tool Database design not addressed and dealt with.  *Not all data uploaded to Apex.* |
|  |  |  |  |  |  |

Feedback:

### ASSIGNMENT MARKING SCHEME

*Student no: Group: Date:*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ***Criteria*** | ***>85%*** | ***Level: 70%+ 1st*** | ***60 – 69% 2:1*** | ***50 – 59% 2:2*** | ***40 – 49% 3rd*** | ***<39% Fail*** | ***<30%*** |
| **Task 3: Design of the ETL and populate star schema**  **(20 marks marks)** | Exceptional work as for first and also showing unique/ self-taught /outstanding application of tools, research, understanding | *All tables fully populated from both “data-bases”. Data Extract, Quality, transformation, calculation of measures fully evidenced using mostly SQL. Competent use of appropriate, advanced SQL (eg sequences etc). Evidence of SQL running successfully.*  *Code is significantly developed from that given in tutorials. Uses standards and includes comments.*  Evidence of SQL running successfully.  *Excellent understanding demonstrated.* | *All tables populated from both “data-bases” as specified. Some Data Extract, Quality, Transformation, calculation of measures evidenced (mostly using SQL). Evidence of SQL running successfully.*  *Code is partially developed from that given in tutorials.*  Evidence of SQL running successfully.  *Good understanding demonstrated.* | *All tables populated from both “data-bases” as specified. Some ETL stages addressed. Maybe errors in code or population. Evidence of SQL running successfully, may be very similar to that given in class.*  *Understanding demonstrated.* | *Some tables populated.*  *Some evidence of SQL running successfully.*  Maybe lack of evidence of the SQL successfully running, or output that omits information.  *Some understanding demonstrated.* | *Little population of db*  *Little evidence of the SQL running successfully.*  *Little understanding demonstrated*. | Work evidences very little understanding in the field of study. Very little done or not relevant to task |
| **Task 4:**   1. **OLAP**   **OR**  **Apex dashboard**  **(15 marks)** | Exceptional work as for first and also showing unique/ self taught /outstanding application of tools, research, understanding. | OLAP: An excellent pivot table produced with very relevant and suitable reports – using appropriate visualisations, labels, titles etc.  An apex dashboard created to meet KPIs, excellent HCI, data consideration, advanced code.  Evidence of research being appropriately applied to drive the areas and types of investigation. Bibl’y  Excellent understanding. | OLAP: A very good pivot table produced with very relevant and suitable reports – using correct visualisations, labels, titles etc.  An apex dashboard created to meet KPIs, some HCI, data consideration, correct code.  Evidence of research some application. Bibliography.  Good understanding. | OLAP: A pivot table produced with some reports –visualisations, labels, titles may not be useful.  An apex dashboard created to meet basic KPIs, some HCI, data consideration, basic SQL.  Some evidence of research – described rather than applied.Bibl’y.  Reasonable understanding demonstrated. | OLAP: A pivot table produced with report –visualisations, labels, titles may not be useful.  An apex dashboard created to meet very basic KPI(s), very basic SQL or code generated by Apex.  Little research or bibliography.  Some understanding demonstrated. | OLAP: little or not useful pivot table produced. Reports lack relevance.  Little or not useful dashboard created to meet very basic KPI(s), very basic SQL or code generated by Apex.  Little bibliography.  Little understanding demonstrated. | Work evidences very little understanding in the field of study. Very little done or not relevant to task |
| **Task 5:**  **DW Approach**  **(15 marks)**  **considerations** | A well structured report demonstrating critical evaluation, explaining the arguments and considerations well and drawing on literature correctly (ie correctly referenced) | As for a first, with a few areas of ambiguity, or missing.  Literature may be more ‘matched’ that used critically. | Some good research and application to the case study. Mostly as discussed taught in class.  Literature included mostly descriptive. | Some relevant research and application to the case study. Mostly as discussed and taught in class, lack of clarity. Literature included mostly descriptive. | Some research and application to the case study. A number of inconsistencies or lack of evidence.  Little literature included. |
| **Portfolio of evidence of work**  **(20 marks)** |  | All uploads submitted on time, completed to an excellent standard | All uploads submitted on time, completed all work | All uploads submitted, not all work completed by time | Some uploads submitted, and mostly completed by time | One upload or late uploads | No uploads |

Feedback:

**DETAILS OF THE REASSESSMENT**

## Reassessment

Reassessment is to ‘re-do’ either or both components as required. A summary sheet listing all changes is required, this can be easily tracked using the Word, ‘tracking’ facility.

Reassessment date: TBA 2021

## Feedback

Formative feedback will be given in tutorials. General feedback will be given approximately a week after the hand-in date and individual feedback upto 3 weeks after the hand-in date.

Formative feedback is feedback on “what you have done already”, this gives you the opportunity to apply the feedback to your final assignment. Students generally find this kind of feedback very useful and gain better marks as a result.

**Feedback schedule**

|  |  |
| --- | --- |
| **Assignment** | **Feedback** |
| Assignment **by Friday December 11th 2020 23:00** | Formative feedback available as indicated in schedule  General feedback provided via the VLE by week 11.  Individual feedback by end of week 11 |

You are always welcome to make an appointment to discuss your feedback.

# Understanding Your Assessment Responsibilities

Please refer to Course Handbook as appropriate

**Mitigation and Extenuating Circumstances**

If you are experiencing problems which are adversely affecting your ability to study (called 'extenuating circumstances'), then you can apply for mitigation. You can find full details of how to apply for mitigation at:

[http://www.leedsbeckett.ac.uk/studenthub/mitigation.htm](http://www.leedsmet.ac.uk/studenthub/mitigation.htm)

**Late Submission**

Without any form of extenuating circumstances, standard penalties apply for late submission of assessed work. These range from 5% to 100% of the possible total mark, depending on the number of days late. Full details (section C1.5.7) of the penalties for late submission of course work are available at:

[http://www.leedsbeckett.ac.uk/about/files/C1\_Assessment\_-\_General\_Provisions.pdf](http://www.leedsmet.ac.uk/about/files/C1_Assessment_-_General_Provisions.pdf)

**Academic Misconduct**

Academic misconduct occurs when you yourself have not done the work that you submit. It may include cheating, plagiarism and other forms of unfair practice. What is and what is not permitted is clearly explained in *The Little Book of Cheating, Plagiarism and Unfair Practice*, available at:   
[http://www.leedsbeckett.ac.uk/studenthub/plagiarism.htm](http://www.leedsmet.ac.uk/studenthub/plagiarism.htm)

The serious consequences of plagiarism and other types of unfair practice are detailed in section C9 of the Academic Regulations at:  
[http://www.leedsbeckett.ac.uk/about/academic-regulations.htm](http://www.leedsmet.ac.uk/about/academic-regulations.htm)