FIT2094 - FIT3171 Databases
S1 / 2019

Assignment 1 - Database Design - AirTnT

1.0. Scenario.

Note: items marked with a **<caret sign^>** *are deliberately omitted and will be discussed in Section 3.1.*

After graduating, you are approached by a startup company which wants to compete with AirBnB in the accommodation and tourism sector. This startup, AirTnT -- short for "AirBnB, tourism and technology" -- intends to provide a more personalised experience compared to AirBnB, while having lower cost than traditional hotels and travel agencies. In the concept of the 'gig economy', AirTnT seeks to engage freelance tour guides and house-sharers.

AirTnT has customers who are identified by a unique customer id. When a customer checks in to an AirTnT facility (i.e. house-share), AirTnT records the customer's first and last names, address, date of birth, current mobile number, one **<X = unique identifier for identity verification^>**, one **<Y = unique identifier for payment details^>**, and emergency contact number (if they are not already on the system). They also record the date and time of check in. The system needs to maintain a record of all check ins for a particular customer. When a customer checks out, the date and time of their departure for this holiday is recorded.

While in AirTnT, customers are located in a house-share. The house-share is identified by a house-share code and name. AirTnT wishes to record the total number of rooms in each house-share and the number of currently available (empty) rooms. Rooms located in a house-share are assigned a room number within that house - thus, for example, each house has a room number 1. The room's phone number (e.g. if a complimentary VoIP phone is provided by the house's owner) and room type are also recorded. Rooms are classified (their room type) as either fixed or open-plan. Not all rooms are supplied with a complimentary phone.

During a customer's check-in, sometimes they may need to be moved from one room to

another, possibly in a different house-share (e.g. if the room is damaged). If this occurs, the date and time the customer is assigned to the new room and/or house-share are recorded (a history of all such assignments during check-in is required). While in AirTnT each customer is assigned one tour guide (identified by a tour guide id) as their main tour guide. A customer's main tour guide may be in charge of many customers. The AirTnT system records each tour guide's first and last names, phone number, and Instagram handle (e.g. @tour_guide). A tour guide may have one or more specialisations (e.g. foodie tours, nature tours, bicycle tours, etc), but not all tour guides who work for AirTnT have a specialisation.

During their check-in, customers are prescribed with activities as part of their tour experience by tour guides. Activities consist of things such as Pub Visits, Bungee Jumping etc.; they also include activities which might be compulsory such as "AirTnT Bicycle Safety Induction" or "Buy AirTnT Travel Insurance". A customer may have activities prescribed by their main tour guide or any other tour guide working for AirTnT. An activity is identified by an activity code. Each activity has a name (such as "Nevis 134m Bungee Jumping") and includes a description of what the activity involves, the time required for the activity and the current standard cost for this activity. When a particular activity is conducted during a customer's check-in, the date and time when the activity is carried out is also recorded. A particular activity is completed before any further activities are run (two of them cannot occur simultaneously). Some activities, such as Bungee Jumping are carried out by other Tourism Providers, more personalised activities may require a tour guide such as AirTnT Safety Induction.

If a tour guide carries out the activity, the tour guide who completes the activity is recorded (the tour guide who completes the activity may be different from the tour guide who prescribes it). Even if a team of tour guides is involved in the activity (e.g. safety induction), only one tour guide (the most senior tour guide in charge) is recorded as completing the activity.

If an activity is performed by a Tourism Provider (e.g. extreme sports facilities), AirTnT does not record the details of the Tourism Provider who completed the activity.

Not all check-ins require an activity to be carried out (e.g. someone just wants to have a 'staycation' and not go anywhere).

Activities may require "extra" items such as stationery, food, USB sticks (for photos), or name badges. Each item held in stock is assigned an item code by AirTnT headquarters. The item description, current stock and price are recorded. For accounting purposes, each item is assigned to a unique cost centre, such as

Food/Beverage, Photography, or Stationery. A cost centre is identified by a cost centre code and has a recorded cost centre title and admin team member's name (this is a person at AirTnT headquarters who is in charge of buying stock!). The quantity of each item used in a particular activity is recorded.

Customers are billed for the cost for the activity itself and also any "extra" items which are used as part of an activity. (e.g. "Bike Ride and Picnic" consists of the activity cost itself and food items). The billed charge is based on the activity/item cost at the date and time of the activity.

AirTnT also records details of its housekeeping staff and their allocation to work in the house-share. At any point in time a housekeeping staff member can only work in (be assigned to) one house-share. A housekeeper is identified by a unique numeric housekeeper id. AirTnT also records the staff member's first and last name, as well as an alphanumeric Police Check Certificate code (if available). The initial date a housekeeper is assigned to work in a house-share is recorded. Housekeepers may be moved between house-shares as staffing requirements change. When a housekeeper finishes an allocation with a particular house-share the date they finished is also recorded. Within these changes, a house-keeper may return to a house-share they previously worked in; if they do so, a new allocation is recorded.

2.0. Sample Documents / Artifacts

AirTnT: Activity Sheet

CUSTOMER DETAILS --

Customer ID	102345
Customer Name	Kanye East
Date/Time of Check In	17/03/2019 16:25
Main Tour Guide	Marge Beauvoir (ID 1298)

ACTIVITY DETAILS --

Activity Code	15509
Activity Name	Social Tennis
Requested By	Usain Federer (ID 2459)
Performed By	
Performed On	17/03/2019 18:15
Charge to Customer	$25.50

Additional Item Charges for this Activity – GST INCLUSIVE

Item Code	Description	Quantity	Total Item Charge
		Total Extra Item Charges =	$0.00

Activity Code	49518
Activity Name	Picnic With A Local Tour Guide
Requested By	Marge Beauvoir (ID 1298)
Performed By	Homer Samson (ID 804)
Performed On	18/03/2019 12:15
Charge to Customer	$5.00

Additional Item Charges for this Activity – GST INCLUSIVE

Item Code	Description	Quantity	Total Item Charge
FNB0010	Vegetarian Meal	1	$9.80
WIN0021	House White Wine	2	$16.00
		Total Extra Item Charges =	$25.80

AirTnT: Housekeeper Assignment Details

⚠ This is an internal document for use by AirTnT Human Resources personnel.
⚠ Privacy Warning: All unauthorised use including duplication is prohibited without prior written consent.

HOUSEKEEPER --

Housekeeper ID	9876
First Name	Manoj
Last Name	Huynh
Contact Phone	0401 000 002
Police Check Certificate	AU/VIC/4554963
Bankruptcy Check Passed?	**Yes** / ~~No~~

HOUSE-SHARE ASSIGNMENTS --

House-Share Code	House-Share Name	Date Assigned	Date Completed
5027	The Vines Carnegie AU	12/6/2016	2/7/2017
5139	Brunswick Happy Home AU	5/7/2017	31/12/2017
9155	Brenno & Friends Darwin AU	5/2/2018	18/6/2018
15332	KLCC Malls Homestay MY	19/7/2018	

3.0. Tasks.

3.1. Compulsory Requirements

Your Details: Please ENSURE your ID, name, Authcate and unit code (e.g. `12345678 - Jack Ma - jma1001 - FIT2094`) are shown on every page of any document you submit. If a document is a multipage document, such as for the normalisation, please also make sure you include page numbers on every page.

Version Control: All working files, as you work on this assignment task, must be stored in **your personal Google Drive** under a **folder** named with your student ID, Authcate, full name, and unit code (as above) and must show a clear history of development. **Teaching staff require access to your Google folder to track and audit your development progress and investigate academic integrity issues.**

<Items marked with the caret ^ sign>: Note that in the original specifications (Section 1.0), two items for the customer, "one **<X = unique identifier for identity verification^>**" and "one **<Y = unique identifier for payment verification^>**" were deliberately omitted! Why? We want you, as future database analysts, to be able to critically think of (and justify!) certain unknown details when faced with a real scenario.

As a compulsory requirement - **you are required to research what attributes (one for X, one for Y) can be used in the case study above.** Hence - in **Assignment 1B** - for each marked^ item, you need to supply the **attribute name in your model.** More importantly, you must justify your decision for each of X and Y with a clear **one-paragraph description/explanation of what they are!** For each item, **the research and decision process (why you chose it as an attribute and why it is a good choice) should be carefully documented with at 2 references per marked^ item.** These write-ups form a part of Item #7 in Assignment 1B (refer Section 4.2).

EXAMPLE: **let's use a bad example which you cannot use in your own submission (you will get no marks if you do use this example).** *You think that storing a 12 digit credit card number is a good idea for "**<Y = unique identifier for payment verification^>**". Therefore, in Assignment 1B, you will include the attribute name 'credit_card_no' in your models (conceptual and logical). Then, you will write a one-paragraph of why storing someone's credit card is a good idea (e.g. 'we need to store credit card numbers so that AirTnT can easily debit/credit the user… this was also used in products such as blah blah… advantages of storing credit cards as opposed to cheques... etc') as well as 2 references (e.g. Coronel & Morris Textbook Case Studies, real-world business cases, research/journal papers etc) each.*

3.2. Moodle Part A Submission: [5 marks]

1. Using LucidChart, prepare an **INITIAL conceptual model** (Entity Relationship Diagram) for AirTnT.
   - For this initial conceptual model, *include what you see as identifiers (keys) for each entity only (other attributes are **not required**) and all relationships.*
   - **Surrogate keys must not be added to this model**. *Connectivity and Participation for all relationships must be shown on the diagram*. Participation must show both minimum and maximum values (using a single line for 1:1 is not acceptable for this unit).

   This initial conceptual model must be submitted to Moodle as Assignment 1 Part A by the date as stipulated in Section 4.0.

   If this submission is not made by this date you will not be able to submit Assignment 1 Part B.

   Your tutor will provide feedback and guidance based on your submitted initial model which should be integrated into your continuing work in Part B.

3.3. Moodle Part B Submission: [100 marks]

2. Perform **normalisation to 3NF** for the data depicted in the sample AirTnT Activity Sheet and AirTnT Housekeeper Assignment Details.

   During normalisation, you must:

   - **Not** add surrogate keys to the normalisation.
   - You **must** include all attributes (you must **not remove** any attribute as derivable)
   - Clearly show UNF, 1NF, 2NF and 3NF.
   - Clearly identify the Primary Key in all relations.
   - Clearly identify the partial and transitive dependencies (if they exist) in all 1NF relations. You may use a dependency diagram or alternative notation (see the normalisation tutorial sample solution for a possible alternative representation).

   Your attribute names as used in your normalisation and those on your conceptual/logical models must be consistent i.e. the same name used on each for the same property.

3. Using LucidChart, prepare a **FULL conceptual model** (Entity Relationship Diagram) for AirTnT.
   - For this FULL conceptual model, *include what you see as identifiers (keys) for each entity, **all required attributes** and all relationships.* This full model will be based on your feedback from your Part A submission, the normalisation above and further reading of the case study. It may be necessary to revise/update this model while developing your logical model in part 4 below.
   - **Surrogate keys must not be added to this model**. *Participation and connectivity for all relationships must be shown on the diagram*.

4. Based on your final full version of your conceptual model, **prepare a logical level design** for the AirTnT database.
   - The logical model must be drawn using the Oracle Data Modeler. The information engineering or Crow's foot notation must be used in drawing the model.
   - All entities depicted must be in 3NF
   - All attributes must be commented in the database.
   - Sequences must be used to generate numeric primary keys and check clauses must be applied to attributes where appropriate.
   - Be sure to include the legend as part of your model.
   - Note that your Google Drive folder (discussed in S3.1) **must clearly indicate your development history with multiple updates as you work on your model. If your model is added to Google Drive without a clear development history - i.e. we cannot trace the actual progress of your development - IT WILL NOT BE MARKED.**

5. **Generate the schema for the database** in Oracle Data Modeler and use the schema to create the database in your Oracle account. The *only* edit you are permitted to carry out to the generated schema file is **to add header comment/s containing your details (remember: ID, name, Authcate and unit code), drop sequence commands and the commands to spool/echo your run of the script.**
   - Capture the output of the schema statements using the spool command.
   - Ensure your script includes drop table and sequence statements at the start of the script.
   - Name the schema file as **airtnt_schema.sql**.

4.0. Submission Requirements

4.1. Assignment 1A:

DUE DATE PER UNIT GUIDE:
Due date: Week 6, Monday 9 am
All times are Melbourne time - i.e. Australian Eastern Daylight Time

The following files are to be submitted:

- A single page pdf file containing your initial version of your conceptual model. Name the file **airtnt_initial_conceptual.pdf**.

- This file must be created via File - Download As - PDF from LucidChart (**do not use screen capture**) and **must be able to be accessed with a development history via Google Drive.** You can create this development history by uploading successive versions to Google Drive as you work on your model.

4.2. Assignment 1B:

DUE DATE PER UNIT GUIDE:
Due date: Week 8, Monday 9 am
All times are Melbourne time - i.e. Australian Eastern Daylight Time

The following files are to be submitted:

1. A single page pdf file containing your final version of your conceptual model. Name the file **airtnt_conceptual.pdf**. This file must be created via File - Download As - PDF from LucidChart (do not use screen capture), again must be able to be accessed with a development history via Google Drive.

2. A pdf document showing your full normalisation of documents A, B and C showing all normal forms (UNF, 1NF, 2NF and 3NF). Name the file **airtnt_normalisation.pdf**

3. A single page pdf file containing the final logical Model you created in Oracle Data Modeller. Name the file **airtnt_logical.pdf.** This pdf must be created via File - Data Modeler - Print Diagram - To PDF File from within SQL Developer, do not use screen capture.

4. A zip file containing your Oracle data modeler project (in zipping these files be sure you include the .dmd file and the folder of the same name). Name the file **airtnt_oraclemodel.zip.**

- This model must be able to be opened by your marker and contain your full model *otherwise your task 4 will not be marked*. For this reason, you should carefully check that your model is complete - you should take your submission archive, copy it to a new temporary folder, extract your submission parts, extract your model and ensure it opens correctly *before* submission.

- *If your marker cannot open your file, as above, task 4 will be zero.*

5. A schema file (CREATE TABLE statements) generated by Oracle Data Modeller. Name the file **airtnt_schema.sql**

6. The output from SQL Developer spool command showing the tables have been created. Name the file **airtnt_schema_output.txt**

7. A pdf document containing any assumptions you have made in developing the model or comments your marker should be aware of. Name the file **airtnt_assumptions.pdf**

Note that there are *seven required files as itemised above*. These files must be zipped into a single zip file named:
a1combined-<ID>-<name>-<authcate>-<unitcode>.zip
e.g., **a1combined-12345678-JackMa-jma1001-FIT2094.zip**
… before the assignment due date/time. ID, name, Authcate and unit code

Submit the **a1combined**… .zip file to Moodle before the due date.

Late submission will incur penalties as outlined in the unit guide.

Appendix

Appendix I: Marking Rubric Part A

	Outstanding (Range D - HD)	Adequate (Range P - C)	Not Adequate (N)
Identified the required Entities (1)	- All/most entities identified. - All/most keys are correctly identified.	- Majority of entities identified. - Majority of keys are correctly identified.	- None or few of entities identified. - None or few of keys are correctly identified
Identified the required Relationships (1)	All/most required relationships identified.	Majority of required relationships identified.	None/few required relationships identified. .
Identified correct Connectivity and Participation (1)	All/Most of depicted relationships Connectivity and Participation correctly identified.	Majority of depicted relationships Connectivity and Participation correctly identified.	None/few of depicted relationships Connectivity and Participation correctly identified.
Able to correctly use the required notation convention and be consistent in its usage. (2)	All notations in the model are consistent and follow the unit's ERD standards.	Most notations in the model are consistent and follow the unit's ERD standards.	Few notations in the model are consistent or follow the unit's ERD standards.

Appendix II: Marking Rubric Part B

	Outstanding (Range D - HD)	Adequate (Range P - C)	Not Adequate (N)
Identify the data requirements to support an organisations operations from the supplied case study and expresses these via a database conceptual model. (50)	All AirTnT operations are supported. ● Required number of entities are present ● All/most required attributes and keys have been captured ● Surrogate keys have not been added ● All/most required relationships have been captured ● All/most required cardinality and participation constraints have been captured ● Good choice of attribute for marked^ ones in business case; as well as clear justification and research.	Some AirTnT operations are not supported. ● Majority of required entities are present ● Majority of required attributes and keys have been captured ● Surrogate keys have not been added ● Majority of required relationships have been captured ● Majority of required cardinality and participation constraints have been captured ● Acceptable choice of attribute for the marked^ ones in business case with some research.	Many of the AirTnT operations are not supported. ● None or few of the required entities are present ● None or few of the required attributes and keys have been captured ● Surrogate keys have been added ● None or few of the required relationships have been captured ● None or few of the required cardinality and participation constraints have been captured ● None or poorly justified attributes for marked^ ones.
Understand and follow a database design methodology. (25)	All/majority of the design processes have been correctly followed: ● All/most Normalisation processes are correct ● Dependency diagrams have been provided and match normalisation. ● ER diagram mapped to logical model with only minor errors/omissions. ● SQL Developer Relational model correctly generated from the logical model ● Sequences have been created to provide numeric primary keys where required	Some of the design processes have been correctly followed: ● Majority of Normalisation processes are correct ● Dependency diagrams have been provided and match normalisation in the majority of situations. ● ER diagram mapped to logical model with only a small number of errors/omissions. ● SQL Developer Relational model correctly generated from the logical model ● Sequences have been created to provide numeric primary keys where required in the majority of situations	Few of the design processes have been correctly followed: ● Significant errors during the Normalisation processes ● Dependency diagrams not provided or have major errors ● ER diagram mapped to logical model with errors/omissions. ● SQL Developer Relational model not correctly generated from the logical model ● Sequences have not been created to provide numeric primary keys where required

Appendix II: Marking Rubric Part B (cont'd)

	Outstanding (Range D - HD)	Adequate (Range P - C)	Not Adequate (N)
Understand and apply the relational model principles into practice. (15)	All relational model principles have been followed: ● All/most entities are in third normal form. ● All/most Primary and Foreign keys are correctly identified. ● All/most data integrity requirements (Entity, Referential, Domain) have been correctly identified.	Most relational model principles have been followed: ● Majority of entities are in third normal form. ● Majority of Primary and Foreign keys are correctly identified. ● Majority of data integrity requirements (Entity, Referential, Domain) have been correctly identified.	Few of the relational model principles have been followed: ● None or few of the entities are in third normal form. ● None or few of the Primary and Foreign keys are correctly identified. ● None or few of the data integrity requirements (Entity, Referential, Domain) have been correctly identified.
Able to generate and modify a schema given a logical model in SQL Developer. (5)	The DDL script was executed without errors.		The DDL script was executed with errors.
Able to correctly use the required notation convention and be consistent in its usage. (5)	All notations in the model are consistent.	Some notations in the model are consistent.	Few notations in the model are consistent.