MIS781

Development of a single source of truth database

Great Keppel Island Resort

Prepared by

Mirna Arivalagan - 220142881
Bonnie McCoy - 219149864
David Petrucci - 220215242

Contents

Student contributions

Student	Section
David Petrucci	1.11 Marketing 2.1 Dataset 1: Customer Table - Marketing Department
Mirna Arivalagan	1.12 Development 2.2 Dataset 2: Accommodation Table - Development Department
Bonnie McCoy	1.13 HR 2.3 Dataset 3: Activity Table - HR Department
Team collaboration	All other sections

1. Introduction

The Great Keppel Island Holiday Resort is a world class ecotourism resort with something for everyone. Accommodation and tours cater to guests of all ages and interests, all within a beautiful location where the southernmost tip of the Great Barrier Reef is right on your doorstep. At Great Keppel Island Holiday Resort, we value the beauty of the natural environment around us and are certified partners with EcoTourism Australia.

Since beginning as a humble eco lodge with just a few cabins the resort has steadily grown over many years and now hosts up to 15k guests a year. As the resort has grown the data requirements have also grown and the resort executives are looking to invest in a new database which can act as the single source of truth for data collected by the business.

The aim of this report is to develop a useful database to be utilised by three key stakeholders at the resort; Marketing, HR and Development. In the first section of this report the objectives, benefits and business rules will be detailed. The second section details the development of a robust database to satisfy the requirements of key stakeholders.

1.1 Objectives for database design

1.11 Marketing

The audience for the customer tables is the marketing department. The data capture and subsequent analysis provided by the customer tables will support the marketing department to maintain and develop the resort's loyalty program. The loyalty program leverages repeat and referral business processes to meet its organisational objectives of increased customer retention and market share. This best practice approach incorporates the following attributes

- Value added customer communications
- Linking rewards to frequency of resort visitation
- Allowing rewards to be transferable to friends and family

Customer categorisation occurs with a tiered loyalty program, with the attained level of status reflected in all customer communication pieces. A customer progresses through the levels of bronze, silver and gold via frequency of resort visitation and is rewarded with the addition of more luxurious room amenities (late checkout, complementary wine and chocolates).

The rewards extend to an offer program that celebrates both the customers 5th stay and 12 month anniversary of their first stay. The gift of 1 night's free accommodation can be used by the customer or transferred to a friend or family member.

Lost customers are tracked with the database recording their last visit date and when it exceeds an 18 month time period, and they don't have any future bookings. This event triggers a communication piece, reminding them it might be time to take a holiday and includes the current offers and packages.

1.12 Development

The development department will use the accommodation and bookings table to understand the most popular accommodation type and develop an ROI report for a new development proposal. Data for the accommodation tables has been extracted from the customer bookings database and summarised by each room that is available at the resort. The analysis will cover if there are certain types of accommodation that are prone to higher cancellations or long term stays to support the decision on the type of accommodation to develop on the resort. It will also cover the costs that need to be considered to run each room type as certain accommodation types has more inclusions and catering that needs to be factored in. The report will also touch on the human resources cost to manage each accommodation type, as certain accommodation types such as the hotel division requires someone with a higher skill set to manage, thus higher level of pay is required.

1.13 Human Resources

The HR department will be utilising the Host table to better understand and optimise scheduling based on pay, break requirements and skill levels. The HR team would like to be able to create dashboards to report on the ratio of capacity to staffing to calculate human resource cost per tour and compare that to the revenue generated by the tour. This will help HR understand where there are requirements for more staffing and where tours are over staffed. They want to be able to offer a variety of time slots to guests while remaining profitable.

1.2 Benefits of the database design

The design of the database conforms to normalisation principles and achieves a level of 3NF. The advantages of utilising normalisation as per (Hoffer, Ramesh, & Topi, 2016) are

- Minimize data redundancy, thereby avoiding anomalies and conserving storage space
- Simplify the enforcement of referential integrity constraints
- Easier data maintenance (insert, update, and delete)

A database is considered to be in 3NF when the following parameters have been met.

Normal Form	Definition
1NF	Each row is unique and all values are atomic
2NF	Tables in 1NF + all non-key attributes are dependant on all of the primary key
3NF	Tables in 2NF + all non-key attributes are dependant only on the primary key

(Carpenter, 2008)

1.3 Assumptions

Table of Assumptions	
1	It assumed that all other operating costs for the investment accommodations such as water/electricity/water and cleaning have previously been factored in and will not form part of the accommodation type investment analysis. Catering costs have also been provided separately which can be used alongside the investment tables to factor in the total cost to run each investment.
2	Manager maximum salary is determined by the Pay Grade level.
3	There will be other tables (e.g. activity bookings, accommodation bookings, etc.) added to this database to complete the single source of truth dataset but their development is not detailed in this report.

1.4 Business Rules used in this report

1. Host and Activities	
1.1	There must be at least one host assigned to each slot but there could also be more than one
1.2	The running cost is related to the pay grade of the host assigned to the activity
1.3	The capacity is the total number of attendees regardless of age
1.4	Hosts cannot be scheduled to more than 3 different activities in one week
1.5	Host salary is exclusive of bonuses, tips and non-monetary work perks
2. Customer	
2.1	First stay date is static and maintained for the life of the customer
2.2	Last stay date is updated to the most recent stay
2.3	Lost customer queries are run utilising the last stay date whereby a duration of 18 months since the last stay and no future bookings triggers the lost customer marketing event
2.4	There are 2 available offers for customers; • The Anniversary offer celebrates 12 months having elapsed from their first stay • The 5th stay offer celebrates the customers 5th stay over their life time as a customer
2.5	Customers are recognised with 3 loyalty tiers • Bronze signifies 4 stays or less • Silver signifies 5-8 stays • Gold signifies 8+ stays
3. Accommodation	
3.1	Successful stay count is an average calculated field based on the booking information from the customers' database. This is the sum of the number of bookings for each room in the calendar year divided by 12 months to obtain the monthly average count.
3.2	Cancellation count is an average calculated field based on the booking information from the customers' database. This is the sum of the number of cancelled bookings for each room in the calendar year divided by 12 months to obtain the monthly average count.

3.3	Cancellation is defined by bookings that are cancelled 7 days or less from the booking start date. Bookings that are cancelled more than 7 days from the booking start date is considered a "lost booking" and is not included as a cancelled booking.
3.4	Long term stays are defined when the difference between the booking start date and booking end date is greater or equals to 7 days.
3.5	Long term stay count is an average calculated field based on the booking information from the customers' database. This is the sum of the number of bookings where the difference between the booking start and end date is greater or equals to 7 days, divided by 12 months to obtain the monthly average count.

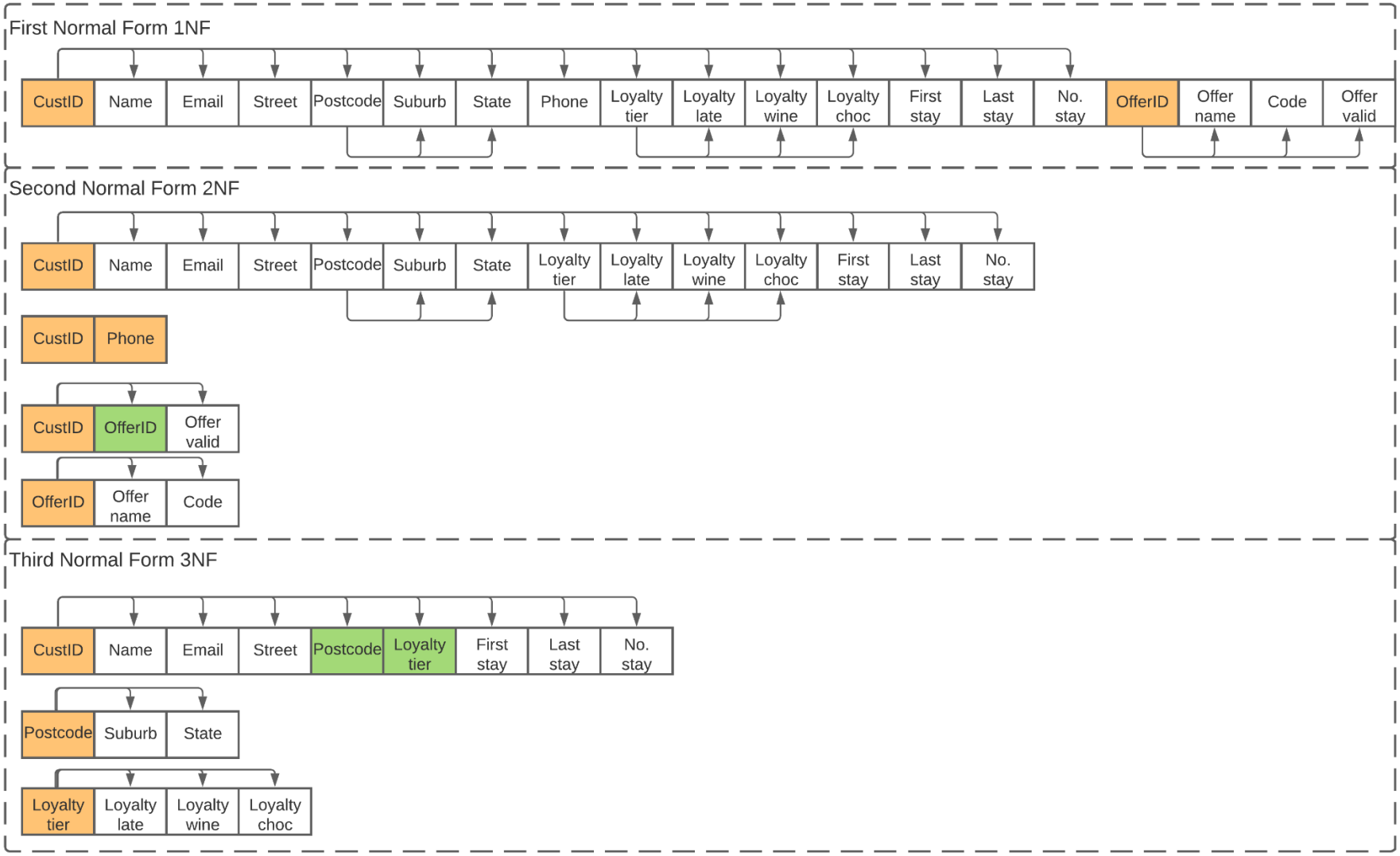
2.1 Dataset 1: Customer Table - Marketing Department

The final Customer dataset is detailed in the table below including; the fields, keys, data types and definitions. In the following sections, the development of the dataset from first normal form to third normal form is detailed.

Table Name	Fields	Data Type	Definition
Customer	CustID(PK)	integer	Unique identifier for the customer
	Name	varchar(50)	Customer first and last name
	Email	varchar(254)	Customer email address
	Street	varchar(50)	Customer street name and type
	Postcode(FK)	integer	Customer postcode
	Loyalty tier(FK)	varchar(6)	Customer loyalty status - gold, silver, bronze
	First stay	date	Date of first customer stay
	Last stay	date	Date of last customer stay
	No. stay	integer	Number of total stays for customer
Phone	CustID(PK)	Integer	Unique identifier for the customer
	Phone(PK)	Varchar(20)	Customer phone number
Customer offer	CustID(PK)	integer	Unique identifier for the customer
	OfferID(FK)	varchar(6)	Unique identifier for the offer
	Offer value	boolean	Offer is applicable to the customer
Offer details	OfferID(PK)	Varchar(6)	Unique identifier for the offer
	Offer name	varchar(50)	Name of the offer
	Code	varchar(25)	Customer redemption code for the offer
Postcode	Postcode(PK)	integer	Customer postcode
	Suburb	varchar(50)	Customer suburb
	State	varchar(50)	Customer Postcode
Loyalty	Loyalty tier(PK)	varchar(6)	Customer loyalty status - gold, silver, bronze
	Loyalty late	boolean	Loyalty tier is valid for late checkout
	Loyalty wine	boolean	Loyalty tier is valid for wine
	Loyalty choc	boolean	Loyalty tier is valid for chocolates

Normalization Summary

The following diagram shows the development of the customer database table from 1NF to 3NF. Details of the transformations and sample data are shown in the following sections.



First Normal Form

All rows are unique and the table only contains atomic values. CustID and OfferID form the composite primary key.

CustID	Name	Email	Street	Postcode	Suburb	State	Phone	Loyalty tier	Loyalty late	Loyalty wine	Loyalty choc	First Stay	Last stay	No. stay	OfferID	Offer name	Code	Offer valid
1	Bob Smith	bsmith@gmail.com	6 Milton Street	3046	Glenroy	VIC	03 9884 8054	Gold	Yes	Yes	Yes	1/11/19	1/06/20	10	ANN	Anniversary	Birthday	Yes
1	Bob Smith	bsmith@gmail.com	6 Milton Street	3046	Glenroy	VIC	0423 122 976	Gold	Yes	Yes	Yes	1/11/19	1/06/20	10	STAY	5th stay	Free stay	Yes
2	Chase Gordon	thechase@gmail.com	8 Aumann Street	3135	Heathmont	VIC	0400 145 876	Bronze	Yes	No	Yes	6/11/19	10/10/20	3	ANN	Anniversary	Birthday	Yes
2	Chase Gordon	thechase@gmail.com	8 Aumann Street	3135	Heathmont	VIC	0400 145 876	Bronze	Yes	No	Yes	6/11/19	10/10/20	3	STAY	5th stay	Free stay	No
3	Christina Lumley	lumleyc@gmail.com	36 Church Street	3012	West Footscray	VIC	0412 554 333	Silver	Yes	Yes	Yes	21/06/20	15/11/20	6	ANN	Anniversary	Birthday	No
3	Christina Lumley	lumleyc@gmail.com	36 Church Street	3012	West Footscray	VIC	0412 554 333	Silver	Yes	Yes	Yes	21/06/20	15/11/20	6	STAY	5th stay	Free stay	No
4	Maria Salvas	m_salvas@optusnet.com.au	4 Gorse Street	2148	Prospect	NSW	02 5562 7615	Gold	No	No	Yes	12/05/20	5/12/20	10	ANN	Anniversary	Birthday	No
4	Maria Salvas	m_salvas@optusnet.com.au	4 Gorse Street	2148	Prospect	NSW	0411 967 544	Gold	No	No	Yes	12/05/20	5/12/20	10	STAY	5th stay	Free stay	Yes
5	Deepti Singh	dsingh_1@gmail.com	24 Bonar Street	4170	Morningside	QLD	0432 988 818	Bronze	No	No	Yes	19/11/20	19/11/20	1	ANN	Anniversary	Birthday	No
5	Deepti Singh	dsingh_1@gmail.com	24 Bonar Street	4170	Morningside	QLD	0432 988 818	Bronze	No	No	Yes	19/11/20	19/11/20	1	STAY	5th stay	Free stay	No

Second Normal Form

Partial dependencies across key fields of CustID and OfferID are separated out into different tables. The Customer Table is comprised of the primary key CustID and non key fields (Name, Email, Street, Postcode, Suburb, State, Loyalty tier, Loyalty late, Loyalty wine, Loyalty choc, First stay, Last stay and No. stay). The CustomerOffer Table is comprised of the composite primary keys of CustID and OfferID and non key field (Offer valid) which is dependant on the CustID. The OfferDetails Table is comprised of the primary key OfferID and the non key fields of (Offer name and code) which are dependant on OfferID .

To ensure all rows remain unique in the Customer Table and to increase database efficiency the phone field is separated into its own table. The Phone Table is comprised of the composite primary keys of CustID and Phone.

Customer Table

CustID	Name	Email	Street	Postcode	Suburb	State	Loyalty tier	Loyalty late	Loyalty wine	Loyalty choc	First Stay	Last stay	No. stay
1	Bob Smith	bsmith@gmail.com	6 Milton Street	3046	Glenroy	VIC	Gold	Yes	Yes	Yes	1/11/19	1/06/20	10
2	Chase Gordon	thechase@gmail.com	8 Aumann Street	3135	Heathmont	VIC	Bronze	Yes	No	Yes	6/11/19	10/10/20	3
3	Christina Lumley	lumleyc@gmail.com	36 Church Street	3012	West Footscray	VIC	Silver	Yes	Yes	Yes	21/06/20	15/11/20	6
4	Maria Salvas	m_salvas@optusnet.com.au	4 Gorse Street	2148	Prospect	NSW	Gold	No	No	Yes	12/05/20	5/12/20	10
5	Deepti Singh	dsingh_1@gmail.com	24 Bonar Street	4170	Morningside	QLD	Bronze	No	No	Yes	19/11/20	19/11/20	1

Phone Table

CustID	Phone
1	03 9884 8054
1	0423 122 976
2	0400 145 876
3	0412 554 333
4	02 5562 7615
4	0411 967 544
5	0432 988 818

CustomerOffer Table

CustID	OfferID	Offer valid
1	ANN	Yes
1	STAY	Yes
2	ANN	Yes
2	STAY	No
3	ANN	No
3	STAY	No
4	ANN	No
4	STAY	Yes
5	ANN	No
5	STAY	No

OfferDetails_Table

OfferID	Offer name	Code
ANN	Anniversary	Birthday
STAY	5th stay	Free stay

Third normal form

Transitive dependencies across non key fields of Postcode and Loyalty tier are removed with separate tables. The Postcode Table is comprised of the primary key Postcode and non key fields (Suburb and State) with Postcode remaining as a foreign key in the Customer Table. The Loyalty Table is comprised of the primary key Loyalty tier and non key fields (loyalty late, loyalty wine, loyalty choc) with Loyalty tier remaining as a foreign key in the Customer Table.

Customer Table

CustID	Name	Email	Street	Postcode	Loyalty tier	First stay	Last stay	No. stay
1	Bob Smith	bsmith@gmail.com	6 Milton Street	3046	Gold	1/11/19	1/06/20	10
2	Chase Gordon	thechase@gmail.com	8 Aumann Street	3135	Bronze	6/11/19	10/10/20	3
3	Christina Lumley	lumleyc@gmail.com	36 Church Street	3012	Silver	21/06/20	15/11/20	6
4	Maria Salvas	m_salvas@optusnet.com.au	4 Gorse Street	2148	Gold	12/05/20	5/12/20	10
5	Deepti Singh	dsingh_1@gmail.com	24 Bonar Street	4170	Bronze	19/11/20	19/11/20	1

Postcode Table

Postcode	Suburb	State
3046	Glenroy	VIC
3135	Heathmont	VIC
3012	West Footscray	VIC
2148	Prospect	NSW
4170	Morningside	QLD

Loyalty Table

Loyalty tier	Loyalty late	Loyalty wine	Loyalty choc
Gold	Yes	Yes	Yes
Silver	Yes	No	Yes
Bronze	No	No	Yes

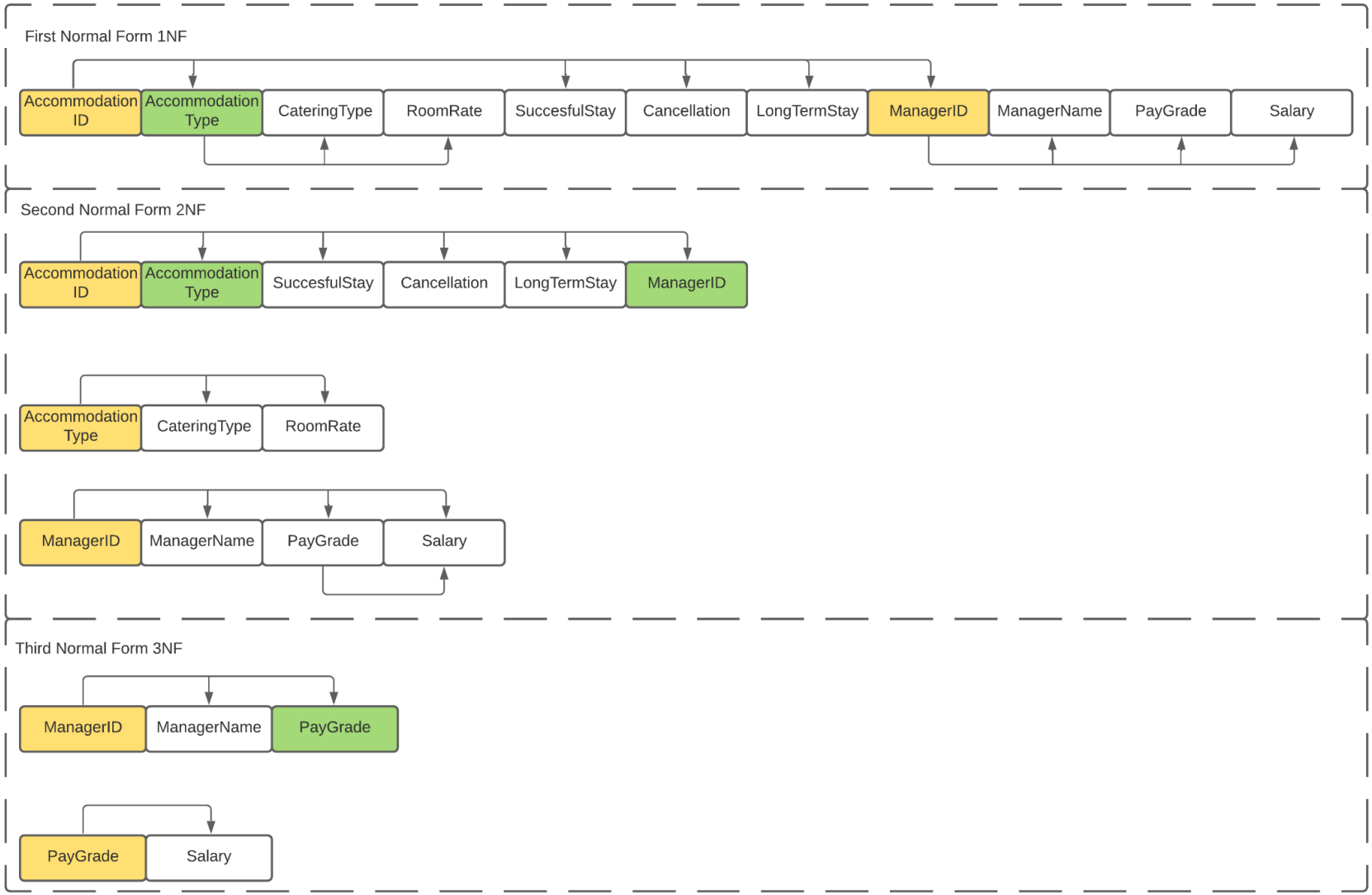
2.2 Dataset 2: Accommodation Table - Development Department

The final Accommodation dataset is detailed in the table below including; the fields, keys, data types and definitions. In the following sections, the development of the dataset from first normal form to third normal form is detailed.

Table Name	Fields	DataType	Definition
Accommodation	AccommodationID(PK)	varchar(15)	Unique ID for each room in Resort
	AccommodationType (PK)	varchar(15)	Name of accommodation type
	CateringType	varchar(15)	Identifier if the room has catering included
	RoomRate	decimal(18,2)	Total cost (AUD) to book room per night
Accommodation Stays	SuccessfulStay	integer	Monthly average of successful stays
	Cancellation	integer	Monthly average of cancellations
	LongTermStay	integer	Monthly average of long term stays
Manager Details	ManagerID (PK)	varchar(15)	Unique identifier of appointed accommodation manager
	ManagerName	varchar(50)	Name of manager
Pay Grade	PayGrade (PK)	varchar(5)	Pay award level
	Salary	integer	The maximum annual salary (AUD)

Normalization Summary

The following diagram shows the development of the accommodation database table from 1NF to 3NF. Details of the transformations and sample data are shown in the following sections.



First Normal Form

This database is in 1NF as all rows are unique and every attribute only contains 1 value. AccommodationID and ManagerID forms the composite key of this database.

Accommodation ID	**Accommodation Type**	Catering Type	Room Rate	Successful Stay	Cancellation	Long Term Stay	**Manager ID**	Manager Name	Pay Grade	Salary
RM001	Hotel Room	Catered	$300.00	25	3	4	FLM001	David Jones	A1	$120,000
GT001	Glamping Tent	Catered	$250.00	17	4	2	FLM002	John Smith	A2	$111,500
CT001	Cottage	Self-Catered	$350.00	24	3	3	FLM003	Jane Doe	B3	$90,000
HV001	Holiday Villa	Self-Catered	$550.00	12	2	4	FLM004	Peter McMillan	A3	110,000
CS001	Camp Site	Self-Catered	$50.00	26	5	1	FLM005	James Finlay	C1	$70,000

Second Normal Form

This table is in 2NF as the partial dependencies between the non key attributes of CateringType, RoomRate and the primary key of AccommodationID has been removed. CateringType and RoomRate is dependent on the AccommodationType, as for each AccommodationType there is more than 1 than 1 AccommodationID, hence why a separate table for the AccommodationType, CateringType and RoomRate was created. A separate table for the ManagerID, ManagerName, PayGrade & Salary has also been created as these are functionally dependent on the primary key of ManagerID.

Accommodation Stays Table

AccommodationID	AccommodationType	SuccessfulStay	Cancellation	LongTermStay	**ManagerID**
RM001	Hotel Room	25	3	4	FLM001
GT001	Glamping Tent	17	4	2	FLM002
CT001	Cottage	24	3	3	FLM003
HV001	Holiday Villa	12	2	4	FLM004
CS001	Camp Site	26	5	1	FLM005

Accommodation Type Table

AccommodationType	CateringType	RoomRate
Hotel Room	Catered	$300.00
Glamping Tent	Catered	$250.00
Cottage	Self Catered	$350.00
Holiday Villa	Self Catered	$550.00
Camping Site	Self Catered	$50.00

Manager Table

ManagerID	ManagerName	PayGrade	Salary
FLM001	David Jones	A1	$120,000
FLM002	John Smith	A2	$111,500
FLM003	Jane Doe	B3	$90,000
FLM004	Peter McMillan	A3	$110,000
FLM005	James Finlay	C1	$70,000

Third Normal Form

This table is now in 3NF as we have removed some transitive dependencies in the tables. Salary was transitively dependent on the ManagerID via the PayGrade. The manager table in 2NF has been split further into 2 separate tables - Manager Table & Pay Grade Table. ManagerID is the primary key and PayGrade is the foreign key in the Manager Table. In the Pay Grade table, PayGrade is now the primary key.

Manager Table

ManagerID	ManagerName	PayGrade
FLM001	David Jones	A1
FLM002	John Smith	A2
FLM003	Jane Doe	B3
FLM004	Peter McMillan	A3
FLM005	James Finlay	C1

Pay Grade Table

PayGrade	Salary
A1	$120,000
A2	$111,500
B3	$90,000
A3	$110,000
C1	$70,000

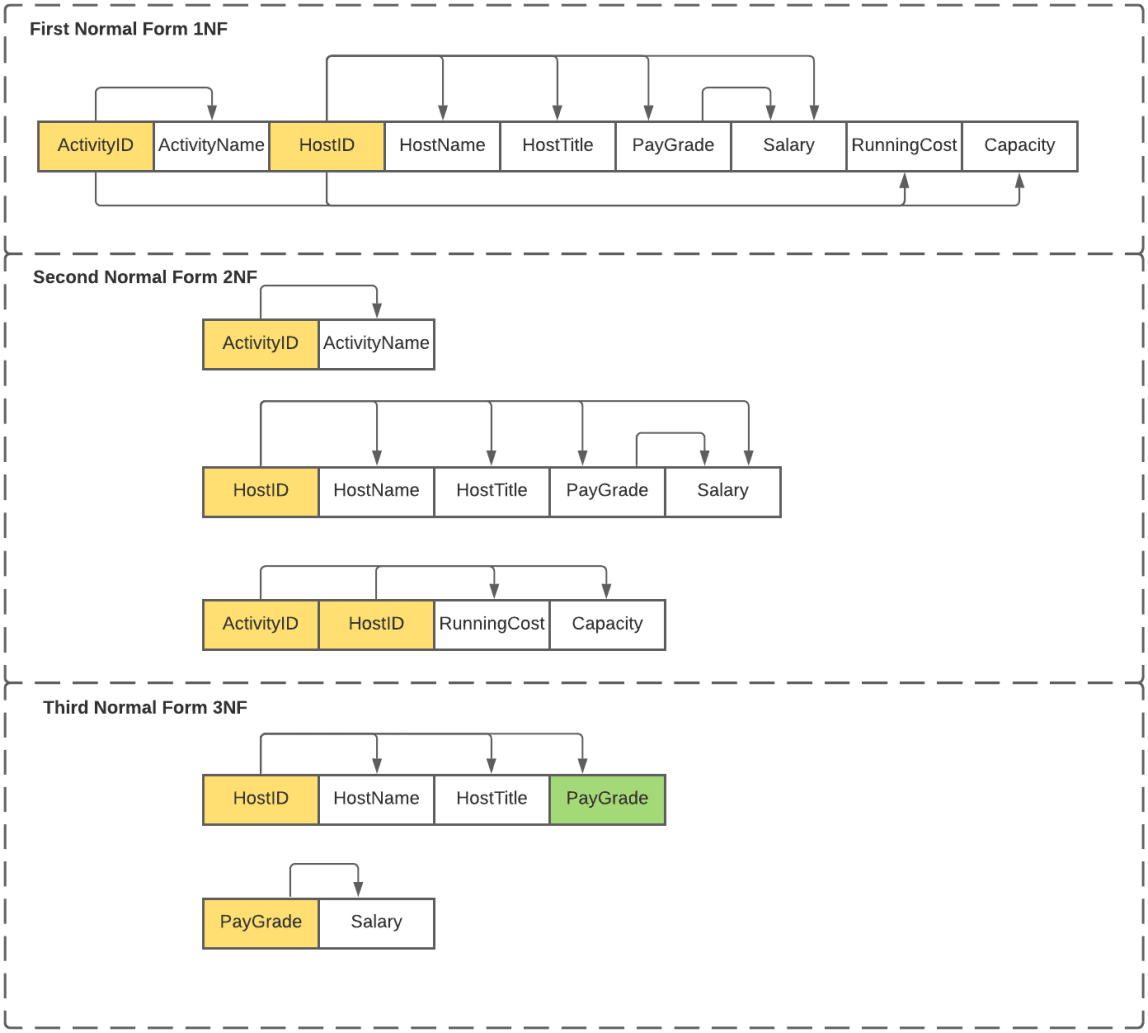
2.3 Dataset 3: Activity Table - HR Department

The final Activity dataset is detailed in the table below including; the fields, keys, data types and definitions. In the following sections, the development of the dataset from first normal form to third normal form is detailed.

Table Name	Fields	DataType	Definition
Activity_Host	ActivityID (PK)	integer	Unique identifier for the Activity
	HostID (PK)	integer	Unique identifier for the Host
	RunningCost	decimal(18,2)	The cost to the business ($AUD) associated with running the tour
	Capacity	integer	The number of attendees that can be accommodated on the activity
Host	HostID (PK)	integer	Unique identifier for the Host
	HostName	varchar(50)	The host's first name and last name
	HostTitle	varchar(150)	The host's Job title
	PayGrade (FK)	varchar(10)	The pay grade that the host title is associated with
PayGrade	PayGrade (PK)	varchar(10)	The pay grade that the host title is associated with
	Salary	int	The annual salary ($AUD) that the host earns
Activity	ActivityID (PK)	integer	Unique identifier for the Activity
	ActivityName	varchar(150)	The name of the activity

Normalization Summary

The following diagram shows the development of the activity database table from 1NF to 3NF. Details of the transformations and sample data are shown in the following sections.



First Normal Form

All rows in the table are unique, no multivalued rows. ActivityID and HostID make up the composite primary key.

ActivityID	ActivityName	**HostID**	HostName	HostTitle	PayGrade	Salary	RunningCost	Capacity
1	Turtle Feeding	1	Nathanael Avery	Lead Animal Feeder	A1	65000	80	20
1	Turtle Feeding	2	Patsy Frost	Apprentice Animal Feeder	C1	35000	60	15
2	Paraglide	3	Keon Lucero	Extreme Sports Expert	A1	65000	95	8
3	Snorkel Adventure	4	Jacqueline Barnett	Marine Generalist	B3	47000	80	14
3	Snorkel Adventure	2	Patsy Frost	Apprentice Animal Feeder	C1	35000	75	12

Second Normal Form

Partial dependencies are removed by splitting the fields determined by HostID (HostName, HostTitle, PayGrade and Salary) into the Host table where HostID is the primary key. Another partial dependency on ActivityID, ActivityName was split out into the Activity table where ActivityID is the primary key. Running Cost and Capacity are functionally dependent on ActivityID and HostID so these are split into the Activity Host table where ActivityID and HostID form the composite primary key.

Host Table

HostID	HostName	HostTitle	PayGrade	Salary
1	Nathanael Avery	Lead Animal Feeder	A1	65000
2	Patsy Frost	Apprentice Animal Feeder	C1	35000
3	Keon Lucero	Extreme Sports Expert	A1	65000
4	Jacqueline Barnett	Marine Generalist	B3	47000

Activity Table

ActivityID	ActivityName
1	Turtle Feeding
2	Paraglide
3	Snorkel Adventure

Activity Host Table

ActivityID	**HostID**	RunningCost	Capacity
1	1	80	20
1	2	60	15
2	3	95	8
3	4	80	14
3	2	75	12

Third Normal Form

Salary is transitively dependent on HostID via pay grade. To remove this transitive dependency and put the tables into 3NF Pay grade and salary were split off from the host table into the salary table where pay grade is the primary key. Pay grade remains in the host table as a foreign key.

Host Table

HostID	HostName	HostTitle	*PayGrade*
1	Nathanael Avery	Lead Animal Feeder	A1
2	Patsy Frost	Apprentice Animal Feeder	C1
3	Keon Lucero	Extreme Sports Expert	A1
4	Jacqueline Barnett	Marine Generalist	B3

Salary Table

PayGrade	Salary
A1	65000
C1	35000
A1	65000
B3	47000

References

Carpenter, D. A. (2008). Clarifying Normalization. *Journal of Information Systems Education*, 379-382.

Hoffer, J. A., Ramesh, V., & Topi, H. (2016). *Modern database management* (12 ed.). Pearson.