Syracuse University
School of Information Studies

Portfolio Milestone Presentation

Richard M. Dzidzornu

05/31/2020

HTTPS://GITHUB.COM/RDZIDZORNU/SU_ADS_PORTFOLIO_MILESTONE

Outline

- Self Introduction

- Professional Background

- Projects Review
  - Portfolio Milestone Overview
  - Information Visualization – IST719
  - Database Admin. & DB Mgt. – IST659
  - Data Analytics – IST707
  - Data Warehousing – IST722

- Learning Objectives

- Next Steps & Conclusion

Self Introduction

Self Introduction

- Born and raised in Ghana, West African.
  - Family
- Relocated into the United Stated
- Hobbies
  - Playing Musical Instruments
  - Soccer
  - Table Tennis(Ping Pong)
- Military Veteran
- Passion for Technology

Professional Background

Professional Background

- University of Ghana Business School (2001 – 2005)
  - BSc. in Business Administration
- Joined the United States Navy 2011 - 2017
  - SQL Training
- University of Maryland University College 2014 - 2016
  - BA in Accounting
- Defense Contract Audit Agency (DCAA) 2017 -2018
  - Auditor
- Deloitte LLC 2018
  - Snr. Consultant

Projects Review

- Portfolio Milestone Overview
- Information Visualization – IST719
- Data Admin.& DB Mgt – IST659
- Data Analytics – IST707
- Data Warehousing – IST722

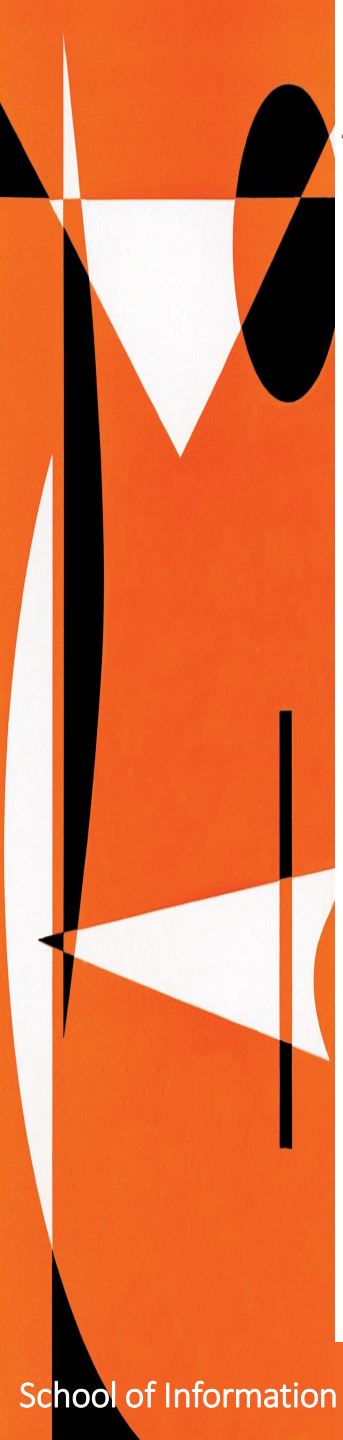
Portfolio Milestone Overview

- Applied Data Science Program at Syracuse University
- The Seven Learning Objectives of Program
  1. Describe a broad overview of the major practice areas in data science.
  2. Collect and organize data.
  3. Identify patterns in data via visualization, statistical analysis, and data mining.
  4. Develop alternative strategies based on the data.
  5. Develop a plan of action to implement the business decisions derived from the analysis.
  6. Demonstrate communication skills regarding data and its analysis for managers, IT professionals, programmers, statisticians and other relevant professionals in their organization.
  7. Synthesize the ethical demonstration of data science practice(e.g., privacy)

Portfolio Milestone Overview

Projects Portfolio Milestone Areas:

- Information Visualization – IST 719
- Data Admin., Concepts & DB Mgt – IST 659
- Data Analytics – IST 707
- Data Warehousing – IST 722



Housing Costs in California Poster

Information Visualization - IST 719

- Information Visualization course was under the instruction and guidance Prof. Gary Krudys.
- Course developed various skills and knowledge in understanding a given dataset, interpret through storytelling and creating powerful visualizations.
- This was translated into the design and development of a poster on Housing Costs in California.
- The goal is to use data collected to create and design, using Adobe Illustrator, a poster that depicts the housing costs in the various districts of California as well factors influencing these costs.

Information Visualization - IST 719

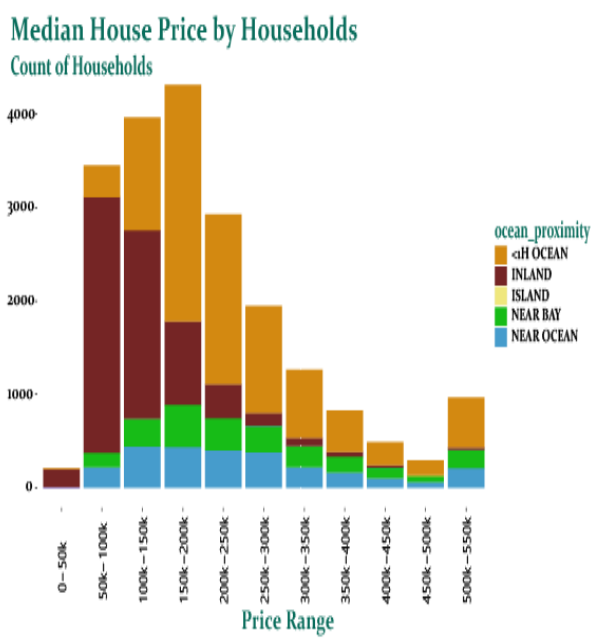
- Dataset drawn from the 1990 U.S Census provides the basis of determining the housing prices in the various districts in California in the poster.

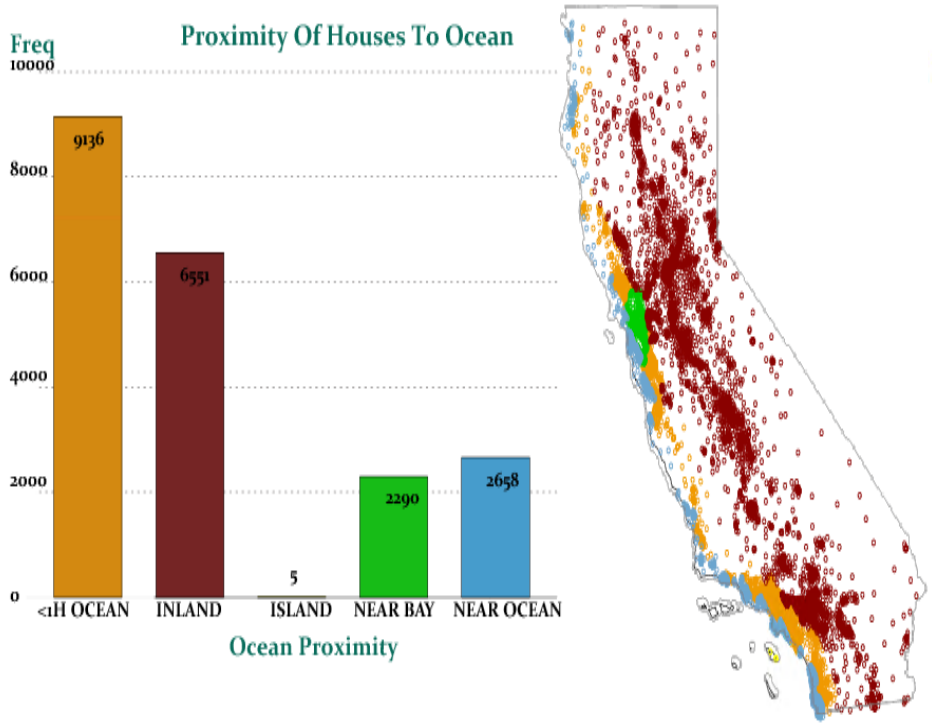
- The dataset has metrics such as the
  - ➔ Population
  - ➔ Median housing price
  - ➔ Latitude
  - ➔ Median income
  - ➔ Median income
  - ➔ Longitude
  - ➔ House median age
  - ➔ Total rooms
  - ➔ Total bedrooms so on for each block group in California.

- R programming language was heavily used for data cleaning, exploration, and analysis.

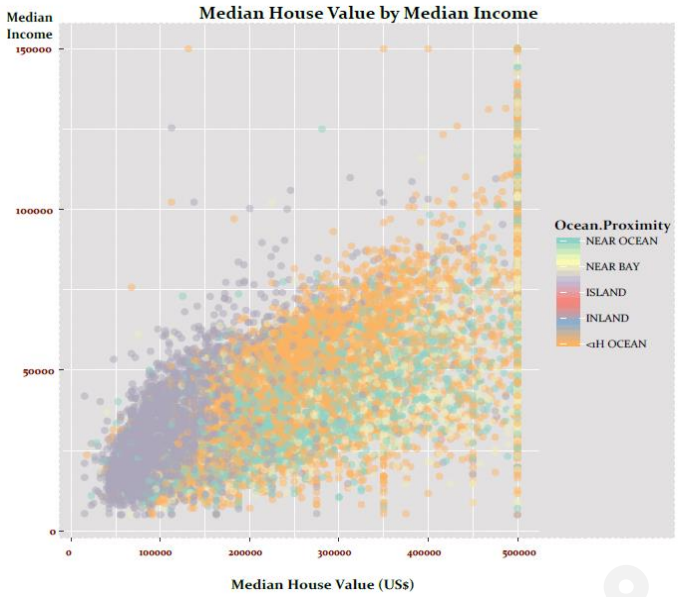
- Initial visualizations were created using various packages in R Studio to answer some of the questions raised in the data exploration process.

Information Visualization - IST 719



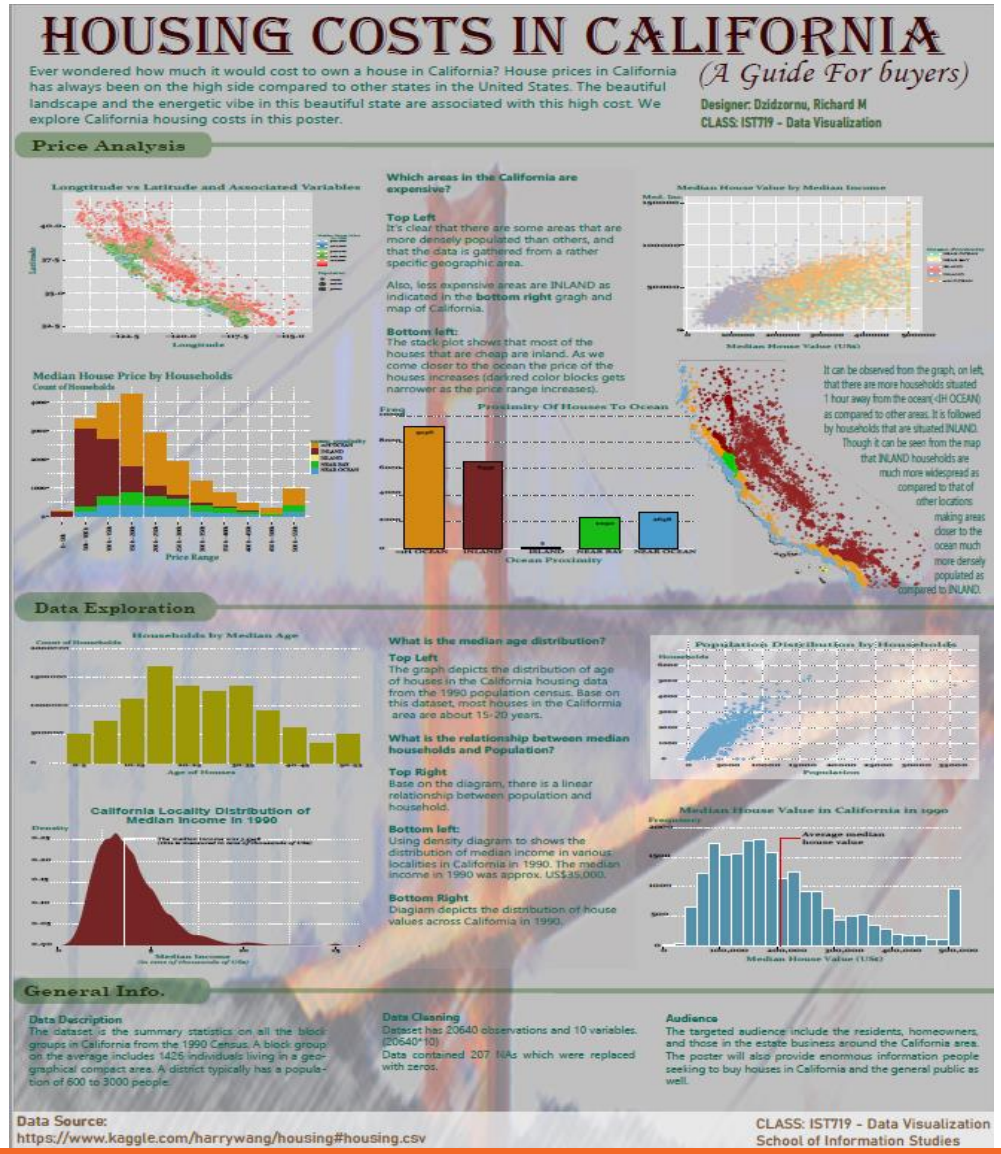


Relationship between median house prices and median Income



Data exploration and analysis of the California housing data covered areas of housing costs in relation to ocean proximity, the population distribution of California from ocean proximity to inland localities

Information Visualization - IST 719



The final piece of the project

Housing Costs in California Poster

The Church Ministry Database

Data Admin. & Database Mgt – IST659

- The Database Administration and Database Management course was taken under the guidance and instructions of Dr. Gregory Block.
- The course provided skills and knowledge in developing and managing different relational databases.
- The Church Ministry Database was developed to organize and track the activities and programs attendance of a local church.
- Tools and applications implements:
  - Microsoft SQL Server
  - Microsoft SQL Server Management Studio
  - Microsoft Visio
  - Microsoft Access
  - Microsoft Excel

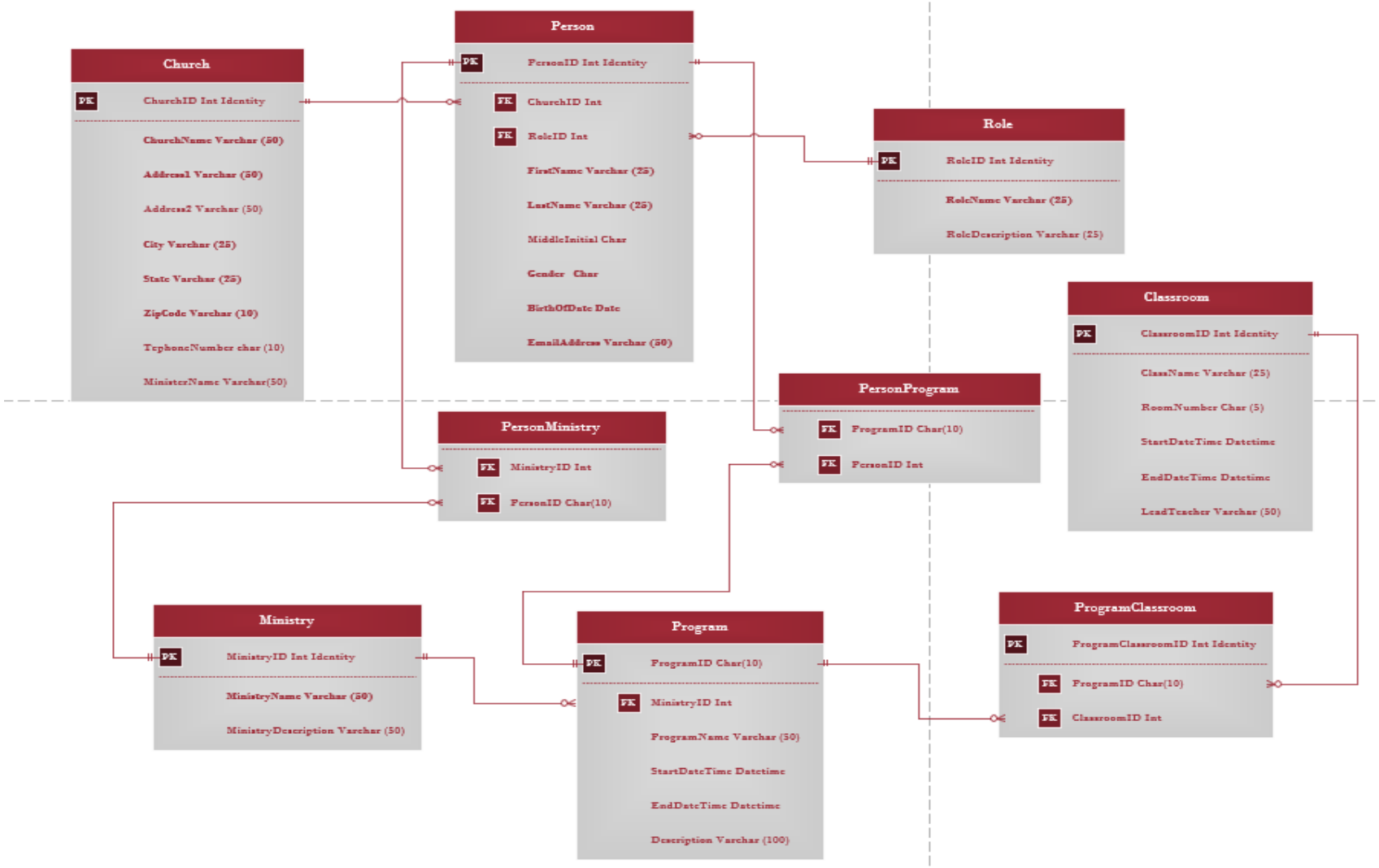
Data Admin. & Database Mgt – IST659

Conceptual Model, DB Objects and Reporting

- The conceptual and physical models were developed using Microsoft Visio
  - ➔ Church
  - ➔ Ministry
  - ➔ Classroom
  - ➔ Person
  - ➔ Program
  - ➔ Role
- Tables, views, functions, stored procedures and other database objects.
  - Microsoft SQL Server application
  - Microsoft SQL Server Management Studio
- User access interface was developed with
  - Microsoft Access
- Interactive reports and visuals/graphs were created using
  - Microsoft Excel

Data Admin. & Database Mgt – IST659

Normalized Logical Data Model



Data Admin. & Database Mgt – IST659

Project.Church Table

Results | Messages

	ChurchID	ChurchName	Address1	Address2	City	State	ZipCode	TelephoneNumber	MinisterName	BranchName
1	1	Visitor	Visitor	Visitor	Visitor	Visitor	Visitor	Visitor	Visitor	Visitor-Visitor
2	2	Qodesh Family Church	14801 Physicians Lane	NULL	Germantown	Maryland	20874	240-410-8457	Pastor Happy Kumah	Qodesh Family Church-Germantown
3	3	Qodesh Family Church	9104 Bakerhill Court	NULL	Gaithersburg	Maryland	20886	240-358-7953	Reverend David Forson	Qodesh Family Church-Gaithersburg
4	4	Qodesh Family Church	5440 Old Tucker Row	NULL	Columbia	Maryland	21044	443-561-9904	Pastor Angel Kumah	Qodesh Family Church-Columbia
5	5	First Love Church	5200 Perring Parkway	Morgan State University	Baltimore City	Maryland	21214	347-621-9159	Reverend Greggory Block	First Love Church-Baltimore City
6	6	First Love Church	2094 North Warwick Avenue	Copping State University	Baltimore	Maryland	21214	410-568-2152	Lady Pastor Sarah Woods	First Love Church-Baltimore
7	7	Qodesh Family Church	7954A Twist Lane	NULL	Springfield	Virginia	22153	703-652-2015	Pastor Darlene Singar	Qodesh Family Church-Springfield
8	8	First Love Church	4400 University Drive	George Mason University	Fairfax	Virginia	22030	615-331-5169	Reverend Anthony Kobi	First Love Church-Fairfax
9	9	First Love Church	1100 Eastern Blvd. N	NULL	Hagerstown	Maryland	21742	258-587-6582	Reverend Edem Ameko	First Love Church-Hagerstown
10	10	Qodesh Family Church	1136 Centerville Tumpike North	NULL	Virginia Beach	Virginia	23320	757-698-1052	Lady Pastor Daniella Ray	Qodesh Family Church-Virginia Beach
11	11	Qodesh Family Church	350 White Horse Avenue	NULL	Trenton	New Jersey	8610	609-556-5854	Pastor William Mensa	Qodesh Family Church-Trenton
12	12	First Love Church	1625 Ocean Avenue	NULL	Brooklyn	New York	11226	347-251-3215	Pastor Jerry Johnson	First Love Church-Brooklyn

Project.Ministry Table

Results | Messages

	MinistryID	MinistryName	MinistryDescription
1	1	Youth Development	Ministry for youth Developing. Age Group 13-17
2	2	Mens	Ministry for Solely Men and is focused on Men Relat...
3	3	Womens	Ministry for Solely Women and focused on Women R...
4	4	Marriage & Family	Ministry for Developing Good Marriages
5	5	Outreach	Ministry for reaching out to others
6	6	Prayer Support	Prayer works Ministry. Pray and Support Others with ...
7	7	Senior Adults	Ministry for Senior Citizens
8	8	Young Adults	Ministry for Individuals in their 20s or Singles
9	9	Divorce Care	Ministry for Supporting People going through Divorce
10	10	Music & Worship	Ministry for Praise and Worship

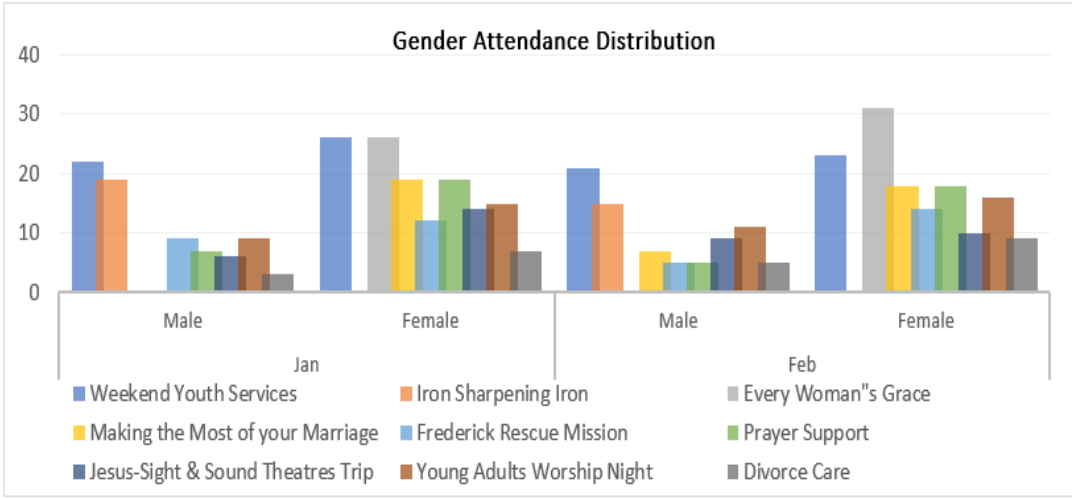
Query executed successfully.

Data Admin. & Database Mgt - IST659

Report

Gender Distribution for the Jan and Feb 2018 Period					
		Jan		Feb	
Ministry	Program	Male	Female	Male	Female
Youth Development	Weekend Youth Services	22	26	21	23
Mens	Iron Sharpening Iron	19	0	15	0
Womens	Every Woman"s Grace	0	26	0	31
Marriage & Family	Making the Most of your Marriage	0	19	7	18
Outreach	Frederick Rescue Mission	9	12	5	14
Prayer Support	Prayer Support	7	19	5	18
Senior Adults	Jesus-Sight & Sound Theatres Trip	6	14	9	10
Young Adults	Young Adults Worship Night	9	15	11	16
Divorce Care	Divorce Care	3	7	5	9

This report basically men are not commited to certain programs or issues



Interactive reports created using Microsoft excel

Data Analytics - IST707

The Patients Reviews on Specific Drugs

Portfolio Review

Data Analytics - IST707

- The Data Analytics course was under the guidance and instructions of Dr. Amy Gates.
- The course introduced various concepts in data mining, data preparation, association rule mining, classification, clustering, evaluation and analysis.
- In the application of these concepts, the Patients Reviews on Specific Drugs project aimed at analyzing the most common conditions for which people takes drugs and the sentiments projected in these reviews.
- R programming language and R Studio were leveraged to conduct this analysis.

Data Analytics - IST707

Initial Data Exploration and Analysis

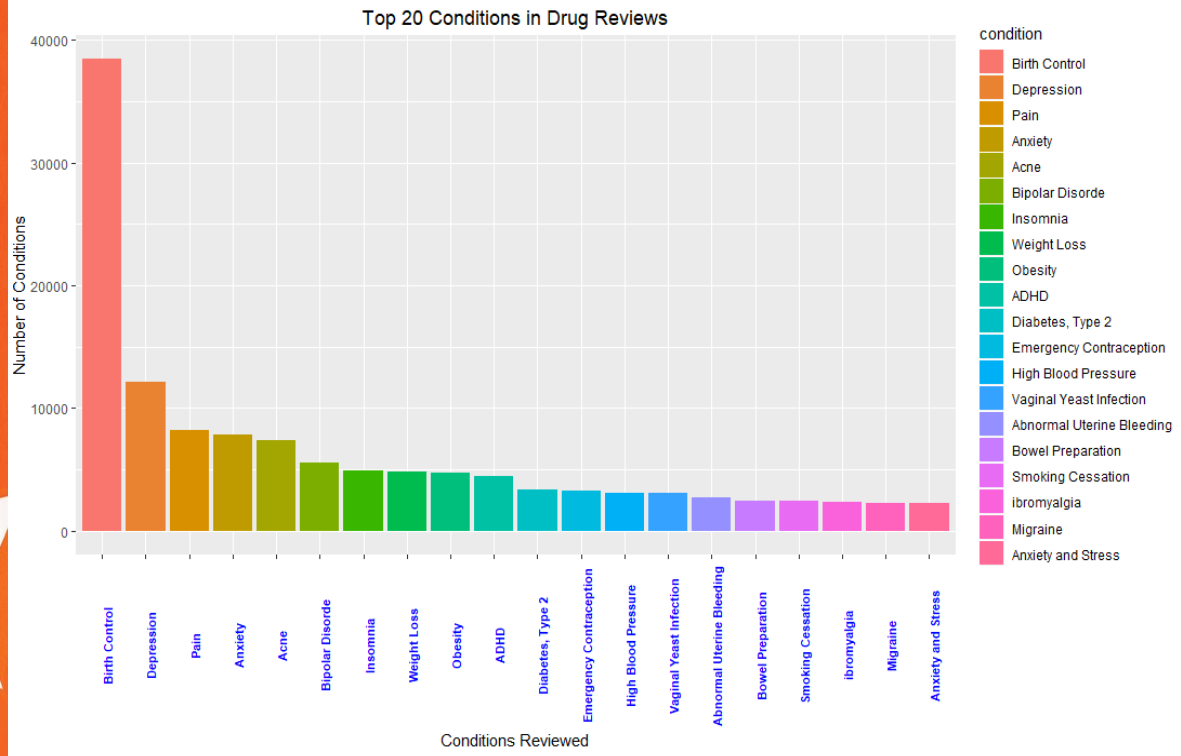
The dataset provided patient reviews on specific drugs along with related conditions and a 10-star patient rating reflecting overall patient satisfaction.

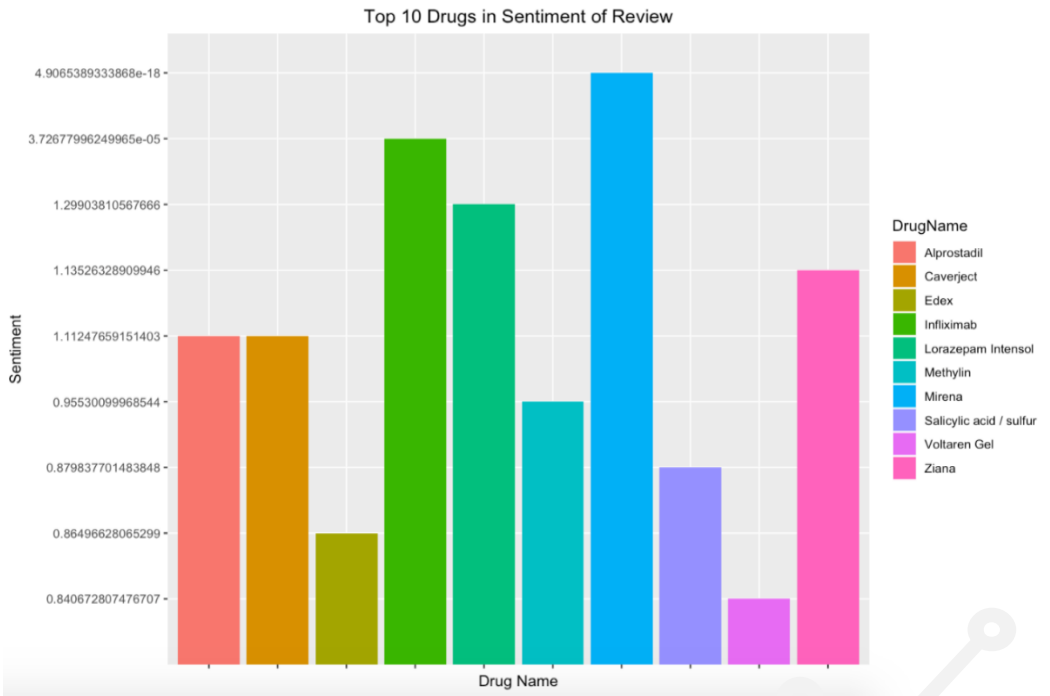
Developed different models to determined correlation between variables such as ratings, conditions, drugs and usefulness of drugs.

The intention was to perform sentiment analysis of drug experience over multiple facets. That is, sentiments on drugs aspects such as effectiveness and side effects.

Data Analytics - IST707

Initial Data Exploration and Analysis

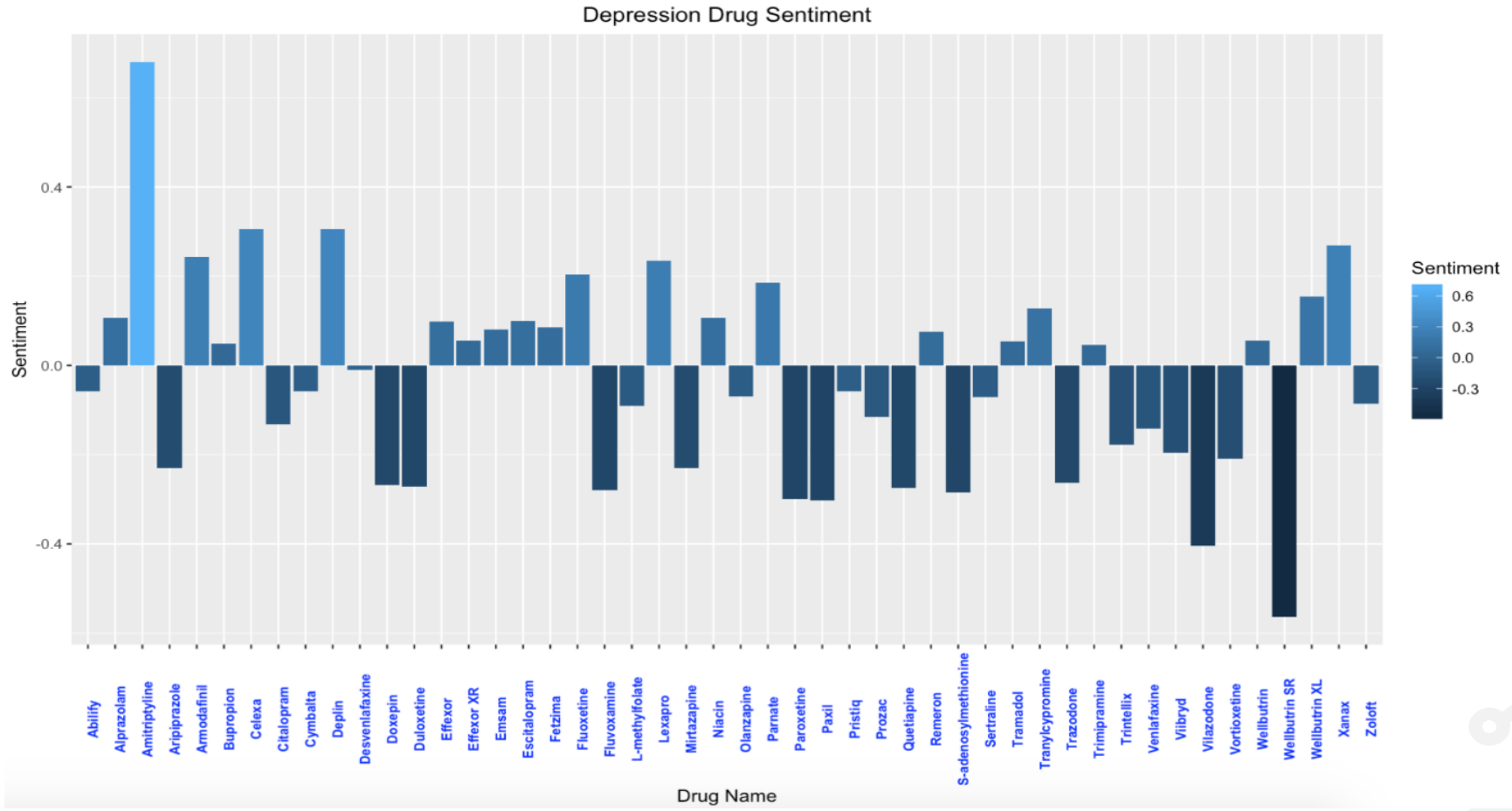




Portfolio Review

Data Analytics - IST707

Sentiment Analysis



Data Warehousing – IST722

Fudgemart Inc.
Data Warehouse

Microsoft SQL Server Integration Services

Data Warehousing – IST722

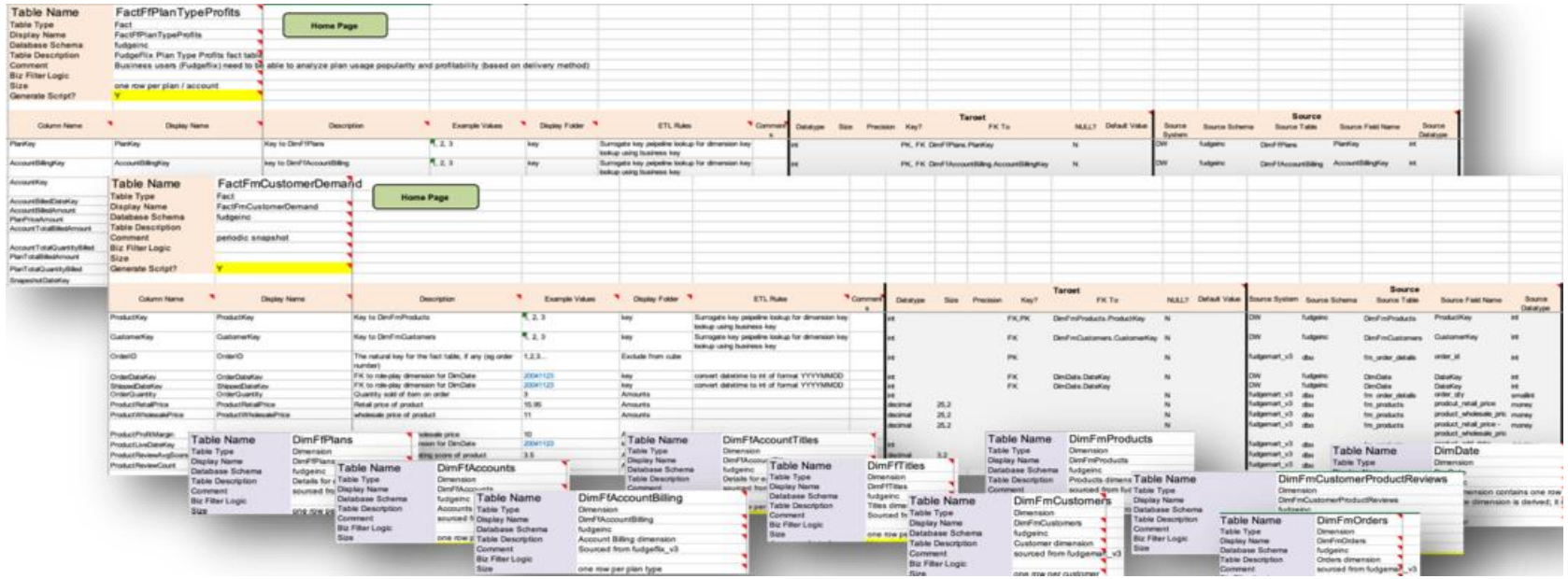
- The Data Warehousing course was under the guidance and instructions of Dr. Gregory Block.
- The course covered concepts, principles, and tools for designing, implementing a data warehouse for Fudgemart Inc. which holds two subsidiaries to store and deliver the organization's data assets which in otherwise was lacking of a centralized data system.
- The data warehouse for Fudgemart Inc. adopted the Kimball methodology or approach to data warehousing.

Portfolio Review

Data Warehousing – IST722

High-level dimensional model

Business Process Name	Fact Table	Fact Grain Type	Granulairty	Facts	FfPlans	FfAccounts	FfAccountBilling	FfAccountBilledDate	FfAccountTitles	FfTitles		FmCustomers	FmProducts	FmOrders	FmCustomerProductRevi	FmOrderDate	FmOrderShippedDate	FmProductReviewDate	FmProductAddDate
Plan Profitability Analysis Business users (Fudgeflix) need to be able to analyze plan usage popularity and profitability (based on delivery method)	FactFfPlanTypeProfits	periodic snapshot	one row per plan / account	plan popularity (i.e., count of customers), plan profitability (i.e., billed amount, not price as the two could be different)	x	x	x	x											
Customer Demand Business users (Fudgeflix) need to be able to analyze revenue by Movie Title and Genre on a weekly, monthly, quarterly and annual baseis	FactFfCustomerDemand	periodic snapshot	one row per movie title, weekly snapshot	average title rating, count of "views", average billed amount, total number of move titles purchased, total number of movie titles by genera purchased			x	x		x									

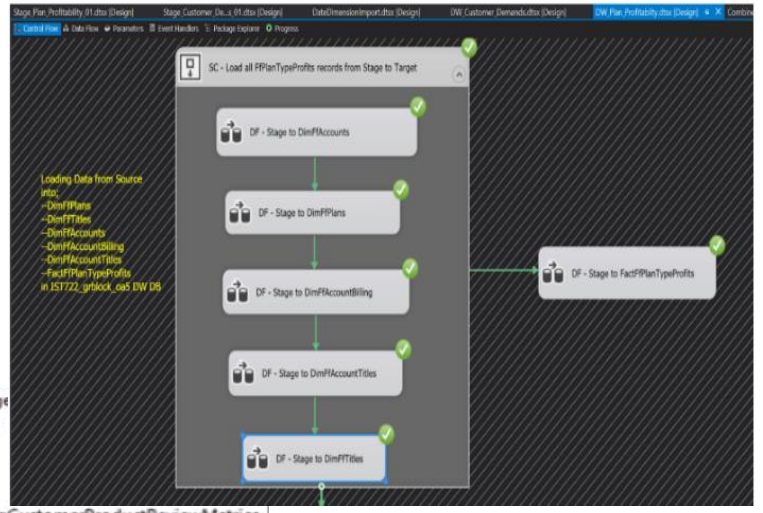


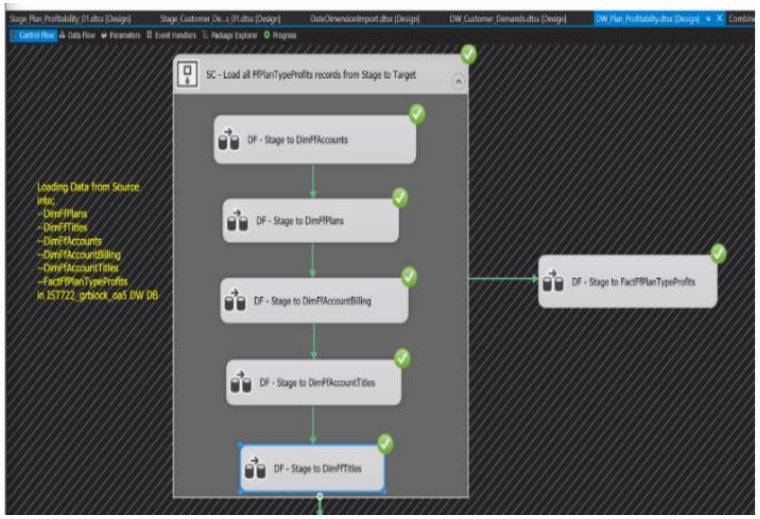
Detailed-level dimensional model

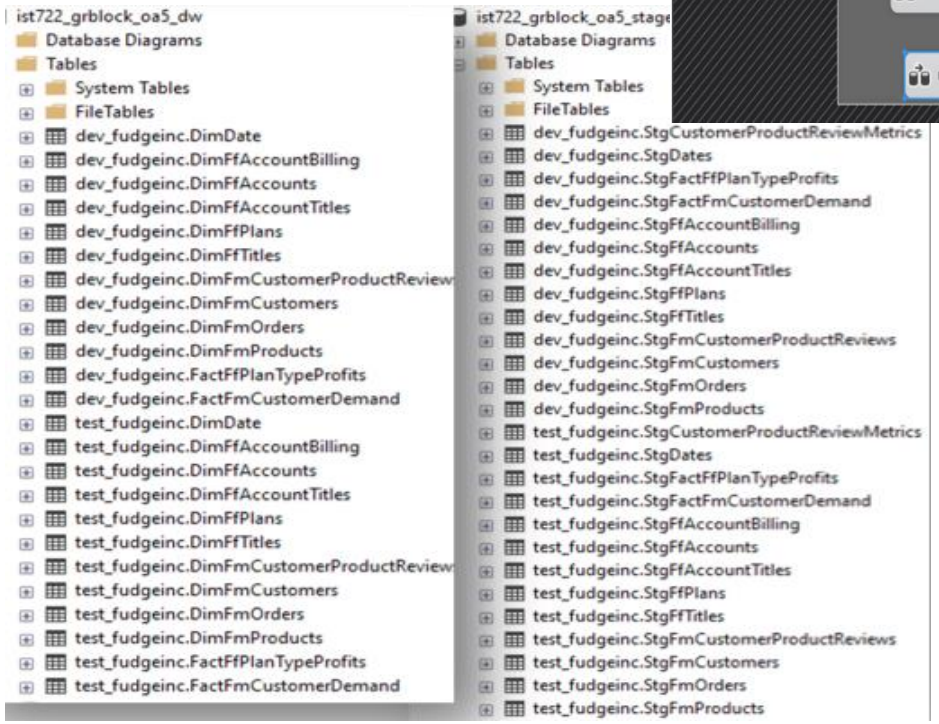
Portfolio Review

Data Warehousing – IST722

Implemented ETL packages in SQL Server Integration Services (SSIS) for source to target mappings







Physical design, the final deliverable to the customer is relational online analytical process (ROLAP) star schema.

The Learning Objectives

The Learning Objectives

- Describe a broad overview of the major practice areas in data

5 types of analytics

- Retrospective
  Traditional business intelligence or reporting
- Diagnostic
  - Analytic dashboards/Drill downs
- Descriptive
  Real time dashboards
- Predictive
  Machine Learning and Forecasting
- Prescriptive
  Make a decision or take action

Data Science

The Learning Objectives

- **Describe a broad overview of the major practice areas in data**

Data Science Requires:

- Statistical Analysis and Programming Skills
- Data Mining
- Regression
- Classification
- Predictive Modeling
- Data Visualization

The Learning Objectives

- Collect and organize data

Data was collected and managed in different ways by implementing database solutions and programming techniques

Data Visualization – IST719: Housing Costs in California, data was derived from the Kaggle and transformed using packages in R programming language and R-Studio

Database Admin – IST659: Data collected and managed using database solutions (MS SQL Server Database and MS Access in creating the Church Ministry Database

Data Analytics – IST707: Gathered data from Kaggle and used Naïve Bayes techniques and sentiment analysis packages in R-studio to organize the data.

The Learning Objectives

- **Identify patterns in data via visualization, statistical analysis, and data mining.**

Information Visualization - IST719: Housing Costs in California
- Statistical analysis and data mining using R programming language
- R packages and Adobe Illustrator for visualizations
- Recommend houses with low costs can be found in the Inlands areas

Data Warehousing – IST722: Fudgemart Inc. Data Warehouse
- Excel and Tableau desktop for both statistical analysis and visualization

Data Analytics – IST707: Patients Review on Specific Drugs
- Statistical analysis and data mining using R programming language
- R packages for visualizations
- sentiment analysis packages in R
- Naïve Bayes techniques for data mining
- Top 5 conditions are birth control, depression, pain, anxiety and Acne

The Learning Objectives

- **Develop a plan of action to implement the business decisions derived from the analysis.**

Data Visualization - IST719: Housing Costs in California

- Noticed Inland population is greater than other areas in terms of proximity to ocean.
- Social amenities should be increased in the Inland
- Businesses in these Inlands will be more viable

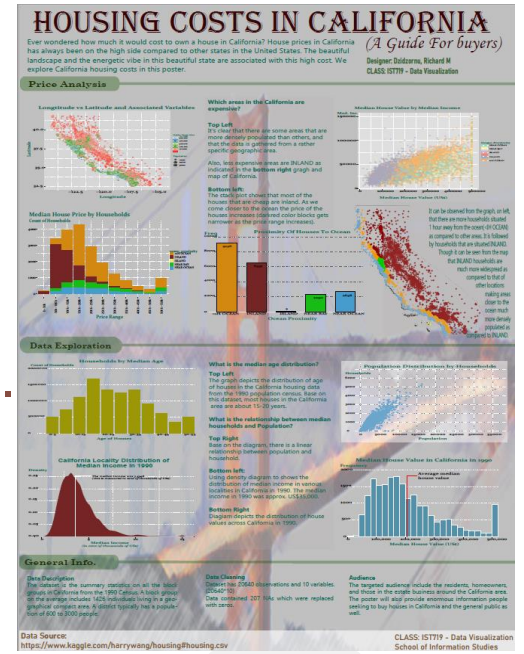
Data Analytics - IST707: Patients Review on Specific Drugs

- More Resources should be put into medical conditions like
  - Anxiety
  - Depression
  - Bipolar Disorders

The Learning Objectives

- **Demonstrate communication skills regarding data and its analysis for managers, IT professionals, programmers, statisticians and other relevant professionals in their organization.**

Communications of results were demonstrated in all PROJECTS



Data Visualization - IST719:

- Results were communicated to the audience in form of a poster
- Presentation of this was made during the final session of this class.
- Recommend affordable houses could be found in the Inlands

Data Analytics - IST707: Patients Review on Specific Drugs

- Results were communicated using graphs generated from the R packages.
- Presentation of this was made during the final session of this class.
- Established the relationship between drugs conditions and sentiments expressed in the reviews

The Learning Objectives

- **Synthesize the ethical demonstration of data science practice(e.g., privacy)**

Ethical issues in data science was demonstrated in all PROJECTS

Data Warehouse IST722: Fudgemart Inc. Data Warehouse

- Dimension Tables with PII were limited to individuals with privy to access such information

Data Analytics - IST707: Patients Review on Specific Drugs

- Exhibiting unbiased representation of results
- Removing details like name and age from data to ensure patients' privacy

Next Step & Conclusion

Next Step & Conclusion

Next Step

Background and experience in Insurance, Auditing, Accounting and Military provide a wide spectrum of layers to bring to the Data Science field.

Continuous education, practice and learn new emerging techniques in the field of Data Science

Become an expert in the areas of ML and AI as the years go by.

Conclusion

The Data Science program from the iSchool provided skills and knowledge for problem solving with multidimensional approach.

The program has undisputedly been a challenge but very rewarding.

The program would highly be recommended hence proud to be a participant of the program.


ThankYou
merci
danke
faleminderit
shukran
terimakasih
dankie
hvala
takk
gràcies
благодаря
ευχαριστώ
kiitos
mahalo
dhanyavād
tak
dank u
děkuji