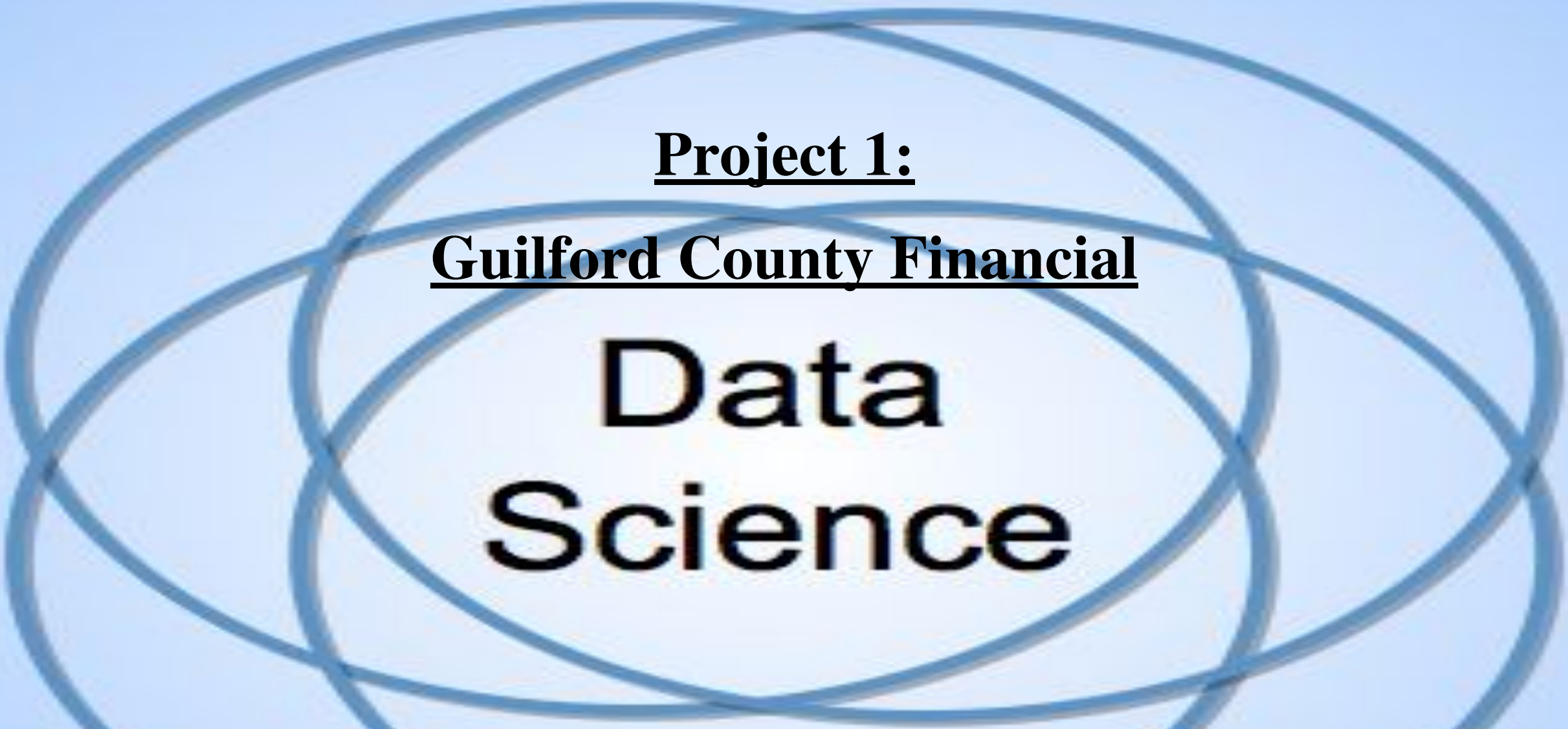
Project 1:

Guilford County Financial



Cody Cothern
Gregory Purvine
Evan Crabtree
Vincent Xiao
Rohit Gulia

Project Overview

What is our dataset?

- Financial Data including adopted/amended budget data and historical transaction data for the Guilford County.

What is our idea?

- Guilford County Financial Modeling based on data for approved budget from 2013 - 2018 and transactions from 2007 - 2018.

What can we do with that idea?

- To help the Financial and Budget Department of Guilford County in maintaining their expected spending and transactions.

Criteria For Success?

- Effective communication/Weekly meetings and each team members being responsible and reliable for individual task(s).

Potential Risks and Obstacles?

- Finding the root cause of the missing transactions.

Goals

- Primary goal: Anomaly detection as the year progresses
  - Applies to individual transactions and total actual spending
- Secondary goal: Reliable projection model
  - Find trends in spending
    - How much is each given entity going to require?
    - What are the general categories for spending?
      - When and where will these purchases be made?

Data Description

- Approved Budgets
  - 2013-2018
  - Departments have sub-departments
  - sub departments perform services
  - expenses are paid from and revenue is paid into funds
  - expenses & revenue come from accounts
  - account represents a specific revenue/expense source
- Transactions
  - 2007-2018
  - transaction occurs between an entity and an account
  - entity - fund & service within government
  - account - expense/revenue source
  - has associated amount, description, and date

Data Dictionary

Files	Field name	Type	Description
Adopted Budgets	Service #	int	6 digit ID number of a service provided by a specific sub dept
	Service Name	String	Name of a service provided by a specific sub dept
	Fund #	int	3 digit ID number of a fund where the money is coming from?
	Fund	String	Name of a fund where the money is coming from?
	Fund- # & Name	String	combines fund number and name
	Dept #	int	3 digit ID number of a dept that provides a category of services
	Dept Name	String	name of a dept that provides a category of services
	SubDept#	int	ID number of a subdept that provides a specific set of services within the dept
	SubDept Name	String	name of a subdept that provides a specific set of services within the dept
	Service Area	String	What type of service does the sub dept provide?
	Type	String	Revenue or expense
	Acct #	int	5 digit ID number of an account where money is either coming from or going to
	Acct Name	String	Name/description of an income source or an expense
	ExpSort		
	Acct Cat	String	Account category: used to group together accounts that are similar
	AU All	String	Service # - Acct #, represents a service - account pair
	budgetType	String	whether or not the budget was adopted
	Amount	int	value of the budget
	Year	int	year the budget was approved for
	PerCapita	double	amount / ? (looks like a magic number at this point)

Data Dictionary

Trans2	Entity	String	Fund#-Service# - corresponds to a fund-service combination in the budget
	Account	int	account # - corresponds to an account number in the budgets spreadsheet
	Date	Date	Date of the transaction
	Description	String	Describes the purpose of the transaction
	Amount	double	amount of the transaction

Approved Budgets Data

Service #	Service Name	Fund #	Fund	Fund- # & Name	Dept #	Dept Name	SubDept#	SubDept Name
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Risk Retention-Liab/Prop/WC	20	Property and Other Ins

Adopted Budgets Data

Service #	Service Name	Fund #	Fund	Fund- # & Name	Dept #	Service Area	Type	Acct #	Acct Name
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	49436	Claims Refunds
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Expense	53270	Claims Paid
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Expense	53120	Vehicle Repair Service
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Expense	53090	Insurance and Bonding
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	45189	Insurance
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	47101	Appropriated Fund Balance
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	49101	Investment Earnings
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	49436	Claims Refunds
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Expense	53270	Claims Paid
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Expense	53120	Vehicle Repair Service
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Expense	53090	Insurance and Bonding
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	45189	Insurance
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	47101	Appropriated Fund Balance
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	49101	Investment Earnings
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	49436	Claims Refunds
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Expense	53270	Claims Paid
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Expense	53120	Vehicle Repair Service
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Expense	53090	Insurance and Bonding
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	45189	Insurance
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	47101	Appropriated Fund Balance
195210	Risk Retention Administration	200	Internal Service	200-Internal Servic	195	Internal Service Fund	Revenue	49101	Investment Earnings

Transaction Data

Entity	Account	Date	Description	Amount
100-100001	41101	7/1/2012	Record PY prpd txs as CY	-116018.63
100-100001	41103	7/1/2012	Record PY prpd txs as CY	-1343129.1
100-100001	41201	7/1/2012	accrue sales&use tax 7/16	1901125.37
100-100001	41202	7/1/2012	accrue sales&use tax 7/16	866051.86
100-100001	41203	7/1/2012	accrue sales&use tax 7/16	575521.51
100-100001	41204	7/1/2012	accrue sales&use tax 7/16	1378.89
100-100001	41212	7/1/2012	accrue sales&use tax 7/16	304822.42
100-100001	41213	7/1/2012	accrue sales&use tax 7/16	631082.97
100-100001	41601	7/1/2012	Accrue Rm Occ & Gr Rec Tx	58746.55
100-100001	42101	7/1/2012	Record PY prpd txs as CY	-6915.34
100-100001	47101	7/1/2012	Reverse PY designated fun	-30838087
100-101110	52050	7/1/2012	AP Accr V067981 P144080	-267.02
100-101110	53040	7/1/2012	AP Accr V107683 PAFP-1301	-194.96
100-101110	53040	7/1/2012	AP Accr V046561 PAFP	-3.27
100-101110	53040	7/1/2012	107683AT&T INC	194.96
100-102110	53040	7/1/2012	AP Accr V107683 PAFP-1301	-48.74
100-102110	53040	7/1/2012	AP Accr V046561 PAFP	-5.04
100-102110	53040	7/1/2012	107683AT&T INC	48.74
100-105002	52010	7/1/2012	AP Accr V031434 P143050	-18
100-105002	52050	7/1/2012	AP Accr V100834 P139734	-10.47
100-105002	53010	7/1/2012	AP Accr V017522 P138415	-900

Related Works

Statsmodels: Econometric and Statistical Modeling with Python

By: Skipper Seabold, Josef Perktold

- Link: file:///C:/Users/cothe/Downloads/Statsmodels_Econometric_and_Statistical_Modeling_w.pdf
- Paper overviewing statsmodel
- statsmodel: a python library used to model statistical data

Data Structures for Statistical Computing in Python

By: Wes McKinney

- Link: https://pdfs.semanticscholar.org/f6da/c1c52d3b07c993fe52513b8964f86e8fe381.pdf
- Paper detailing working with financial data sets.
- Includes use of pandas

Financial Modelling in Python

By: S. Fletcher & C. Gardner

- Link: https://leseprobe.buch.de/images-adb/4c/62/4c6218a0-1d7c-48b7-becf-dc7ddd1b6fa3.pdf
- Book explaining and showing how to use python to model financial data

Building a Financial Model with Pandas

By: Chris Moffit

- Link: http://pbpython.com/amortization-model.html
- Article/blog detailing how to use Pandas to financial models
  - Much more detailed than needed, but touches on some really valuable topics

Quora discussion on detecting anomalies in financial data

By: Various

- Link: https://www.quora.com/What-are-the-best-statistical-methods-for-detecting-anomaly-in-financial-time-series-data
- A very insightful discussion on the various methods for detecting anomalies in financial time series data

Introduction to Anomaly Detection

By: Pramit Choudhary

- Link: https://www.datascience.com/blog/python-anomaly-detection
- A cursory overview of using Python to detect anomalies in financial time series data and the methodologies used
- Also briefly touches on using machine learning for this task

Tasks

- Maintaining Data Dictionary and documentation. → Greg
- Finding the patterns in data. → Rohit
- Which department spends the most and where? → Rohit
- Finding the relationships in the data. → Evan

Tasks

- Find more related works. → Evan
- How the spending trends change with time. → Vincent
- Finding the anomalies in the data. → Cody

Thank you