



ATLAS Analysis

Project Plan

David Gayman
davidrgayman@gmail.com

2

Goals & Tasks

TODO

- Review:
 - Tasks from paper notes
 - Tasks from sticky notes
 - Tasks from repository files
 - Project instructions

2.1

Timeline

Stage	Name	Description
I	Setup	Implement the basic data import and infrastructure.
II	Analysis	Multiple Goals can be worked on at the same time by different people.
III	Conclusions	Draw conclusions and merge into final report and dashboard.

2.2

Key

Priority	Meaning
1	High
2	Medium
3	Low

Normal item
Critical-path item

Priority

Color

2.3

Goal & Task Data

Category	Goal #	Goal Priority	Goal	Task #	Task Priority	Deadline	Task	Assignee
Data Ingestion	G010		Required raw data files collected in repository.	T010		11/5/2020	Configure repository.	
				T020		11/5/2020	Get sales data into raw data directory.	
				T030		11/5/2020	Get marketing communications/attempts/responses into raw data directory.	
				T040		11/5/2020	Get lists of schools that are customers into raw data directory.	

			T050	11/5/2020	Get List of all schools that are possible customers into raw data directory.	
			T060	11/5/2020	Get school locations into raw data directory.	
	G020	Data ingestion implemented.	T070	11/5/2020	Read raw files into DataFrames.	
			T080	11/7/2020	Perform basic cleanup (nulls, invalid characters, etc.).	
			T090		Perform advanced cleanup (collect/combine/pre-process).	
Data Processing	G030	Infrastructure implemented.	T1001	11/7/2020	Implement infrastructure to provide data access to all developers.	David
			T1101		Implement infrastructure to provide collection and handling of processing artifacts (processed data, metrics, and plots -> multiple dashboards, compiled reports).	David
Basic Data Analysis	G040	Analysis procedure and plots implemented.	T1201	11/10/2020	Implement analysis procedure and plots.	
			T1302		Plot trends showing comparison to other educational services. {}	
	G050	Conclusions documented in the final report.	T1401		Answer in the report: "Which quarters, years etc. have the highest sales, and highest growth?"	
			T1502		Answer in the report: "What conclusions can be reached by analyzing the trends in comparison to other educational services?"	
			T1601		Answer in the report: "What actions can be taken to increase sales? (suggestions: Send more/better targeted emails, phone calls, target certain types of customers, would this be applicable to ...)"	
Conversion Rate Analysis	G060	Analysis implemented.	T170		Document the sales process (initial contact, follow-up, negotiation, purchase). {}	
			T180		Compute time differences within stages of the sales process. {}	
			T190		Compute the initial purchase conversion rates. {}	
			T200		Compute the upgrade conversion rates. {}	
	G070	Conclusions documented in the final report.	T2101		Answer in the report: "What is the length of time between initial contact and purchase decision?"	Ibrahima & David
			T2201		Answer in the report: "What is the initial purchase conversion rate based on marketing/sales communications?"	Milad & Jonathan
			T2301		Answer in the report: "What is the upgrade conversion rate (from 3-months to 6-months to 1-year)?"	Milad & Jonathan
Sales Forecasting Analysis	G080	Analysis implemented.	T240		Implement Monte-carlo-based sales forecasting procedure and plots. {}	Milad & Jonathan
			T250		Implement legacy sales forecasting procedure and plots. {}	Milad & Jacinta
	G090	Conclusions documented in the final report.	T2601		Answer in the report: "Taking into account both legacy and Monte-Carlo analyses, what sales are predicted for each quarter through end of 2021?"	Milad & Jonathan
Customer & Market Analysis	G100	Analysis implemented.	T270		Categorize school data (hint: DataFrame.groupby()). {}	
			T280		Plot customers on a map.	
			T290		Analyze data to identify highest-responding, and highest-paying customers.	
			T300		Correlate conversion rates with the customers ranked by highest-responding/highest-paying.	
	G110	Document conclusions in the final report.	T3102		Answer in the report: "What market penetration does the company have into each group of schools (ivy league, state schools, community college, high schools, etc.)"	
			T3201		Answer in the report: "What market penetration does the company have regionally?"	
			T3303		Answer in the report: "Is the company targeting the highest-responding customers?"	
Deliverables	G120	Final Report	T340		Identify and describe the conclusions of the analysis.	All
			T350		Describe the specific analysis procedure used to support the analysis.	All
			T360		Document the specific data sources, resources, links, etc. that were used for analysis, in the appendix.	All
			T370	11/16/2020	Provide report introduction, table of contents, conclusion, and appendix in the deliverable zip file.	Milad
	G130	Dashboard	T3801		Implement a dashboard structure to organize the analysis results and data views.	
			T3901		Implement analysis-specific dashboard views containing analysis plots.	All
			T4002		Implement college-specific dashboard view(s) that display analysis plots for the college in the selected drop-down.	
			T4103		Implement college-specific dashboard view(s) that display pre-processed (unanalyzed) data plots for the college in the selected drop-down.	

			T4202		Implement served/hosted dashboard?	
			T4302		Document how to access the hosted dashboard (if applicable).	
			T4401		Package the dashboard solution in the deliverable zip file.	
	G140	Source code	T4502		Package the source code in the deliverable zip file.	
	G150	Processing artifacts	T4603		Implement data dumps for all critical plots.	
			T4702		Package the important processing artifacts in the deliverable zip file. Examples include key metrics, statistics, data dumps, etc.	