



KIPP Texas College Match & Improving College Persistence

PROJECT PLAN



TCOM 5340 FALL 2020

Carmella Blackwell, Taylor Gilfillan, Gwladys Peg, Melvin Zaldivar

Table of Contents

Section I: Problem Statement	1
Section II: Project Mission Statement	1
Section III: Project Objection	1
Section IV: Project Work Requirements	2
Section V: Work Breakdown Structure	3
Section VI: Exit Criteria	9
Section VII: Milestones	10
Section VIII: Working Schedules.....	11
Section IX: Required Resources	17
Section X: Major Contributors	18
Section XI: Control Systems	23
Section XII: Risk Table	26

I. Problem Statement

The Knowledge Is Power Program (KIPP) Texas Public School district is a collection of 59 schools spread across 4 cities in Texas - Austin, Dallas, Houston, and San Antonio. In total, KIPP Texas serves nearly 30,000 students, with each city/region operating separately. KIPP Texas has a team called the KIPP Through College (KTC) which uses a college match strategy that has been a key component in assisting with college admission and completion for its students. One of the metrics they use when guiding students is called the ECC rate (estimated college completion), which is based on college graduation rates for each school. For example, an ECC rate of 69% means that 69% of students graduated from the school for a recent window of time. This number is also further disaggregated for minorities, which make up more than 98% of students served by the district.

In KIPP Texas - Houston, there is a nearly 30% gap between the number of students going to college and completing college. The KTC team at KIPP Texas-Houston would like to further increase its college completion rates in efforts to ensure all students are able to live choice-filled lives, a promise that cannot wait. This project will analyze the college match strategy to identify areas of improvement so that even more KIPP students can succeed in becoming college graduates in the future.

II. Project Mission Statement

The project team's mission is to develop a written data report analyzing the current and former KIPP student academic profiles and college match data to provide insight and suggestions around ways to decrease the future college drop-out rate for current juniors and seniors.

The goals for this project include:

1. Answer four (4) research questions posed by the KTC team through statistical analysis in R and interactive Tableau visualizations.
2. Communicate key insights and recommendations in a written report to share with internal stakeholders.

III. Project Objective

The objective for this project is to:

Decrease the college dropout rate for freshmen and sophomore college students by 12% (currently high school juniors and seniors) for KTX-Houston alumni by 2023.

IV. Project Work Requirements

To successfully deliver this project, the following work requirements must be met:

1. Written Report

- a. This will be a summary of all findings from four (4) research questions posed by the KTC team.
- b. The report will be comprised of the following section: executive summary, introduction, methods, results, and discussion.
- c. The audience will be the KIPP Texas Strategy Team, and the KIPP Through College team.

2. Tableau Visualization of Data

- a. This will be an interactive way to view and interrogate the takeaways from the research questions posed in four (4) dashboards within one Tableau workbook (4 questions, 4 tabs)
- b. Permissions / Student Data Security requirements:
 - i. Tableau workbook will be shared via SFTP to KIPP Texas to connect through their secure Tableau Server to share internally
 - ii. UHD presentation - Will be presented through screen-sharing only for presentation for student data confidentiality reasons.

V. Work Breakdown Structure

Work Breakdown Structure			
A. Clean Data			
	A.1. Import each data source into R (3 Alumni and 3 High School datasets):		
		A.1.1. Check for completeness and relevance of data	
			A.1.1.1. Remove rows or columns with missing values
			A.1.1.2. Remove columns or features that are not relevant to project
		A.1.2. Check for accuracy of data	
			A.1.2.1. Remove any rows or observations that contain duplicates or errors
		A.1.3. Format structure and set data types	
			A.1.3.1. Ensure data is imported and labeled correctly (e.g. numeric data is labeled as numeric) and convert character data to factors
			A.1.3.2. Separate any data in columns that contain more than one feature and create new feature column(s) when necessary
	A.2. Document the data cleaning process		
	A.3. Write cleaned datasets into files		
	A.4. Get approval from PM for cleaned datasets		
B. Research Question Analysis			
	B.1. Research Question 1: Are there better college retention rates or ECC for in-state / out of state?		
		B.1.1. Research previous analysis of ECC vs college state residency	
		B.1.2. Import appropriate cleaned data set(s) into R and select target features	

B. Research Question Analysis (Continued)		
		B.1.3. Choose and perform the analysis method(s)
		B.1.3.1 Visualize data and choose analysis method(s)
		B.1.3.2. Make connections to the college match strategy (Prepare for Match)
		B.1.3.3. Document the data analysis process
		B.1.4. Write Executive Summary for research question 1
		B.1.4.1. Choose 1-3 key insights gleaned from analysis
		B.2. Research Question 2: Are our kids “beating” the ECC rates? If so, are the factors noted?
		B.2.1. Research any previous analysis of ECC and actual complete rate for colleges
		B.2.2. Import appropriate cleaned data set(s) into R and select target features
		B.2.3. Choose and perform the analysis methods
		B.2.3.1 Visualize data and choose analysis method(s)
		B.2.3.2. Make connections to the college match strategy (Make A Strong Transition)
		B.2.3.3. Document the data analysis process
		B.2.4. Write Executive Summary for research question 2
		B.2.4.1. Choose 1-3 key insights
		B.3 Research Question 3: Is there a correlation between ECC and applying to 6+ schools, or 9+ schools?
		B.3.1. Research previous analysis of number of applications submitted and college completion rates
		B.3.2. Import appropriate cleaned data set(s) into R and select target features

B. Research Question Analysis (Continued)		
		B.3.3. Choose and perform the analysis methods
		B.3.3.1 Visualize data and choose analysis method(s)
		B.3.3.2. Make connections to the college match strategy (Apply to all wish-list colleges)
		B.3.3.3. Document the data analysis process
		B.3.4. Write Executive Summary for research question 3
		B.3.4.1. Choose 1-3 key insights
		B.4. Research Question 4: What is correlation between college ready indicators (TSI, ACT, GPA, AP, 1st year GPA / retention rates, # of applications submitted) and college success? Is there any difference in 2-year and 4-year program success?
		B.4.1. Research previous analysis of number of applications submitted and college completion rates
		B.4.2. Import appropriate cleaned data set(s) into R and select target features
		B.4.3. Choose and perform the analysis methods
		B.4.3.1 Visualize data and choose analysis method(s)
		B.4.3.2. Make connections to the college match strategy (Apply to all wish-list colleges)
		B.4.3.3. Document the data analysis process
		B.4.4. Write Executive Summary for research question 4
		B.4.4.1. Choose 1-3 key insights
D. Tableau Visualization		
		D.1. Connect Data
		D.1.1. Import each dataset into Tableau (separate data sources in same workbook)

D. Tableau Visualization (Continued)		
		D.1.2. Label Data sources by research question (RC1, RC2, RC3, RC4)
	D.2. RQ1 Visualization	
	D.2.1. Explore data visually	
		D.2.1.1. Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)
	D.2.2. Organize visual(s) into dashboard view	
	D.3. RQ2 Visualization	
	D.3.1. Explore data visually	
		D.3.1.1. Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)
	D.3.2. Organize visual(s) into dashboard view	
	D.4. RQ3 Visualization	
	D.4.1. Explore data visually	
		D.4.1.1. Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)
	D.4.2. Organize visual(s) into dashboard view	
	D.5. RQ4 Visualization	
	D.5.1. Explore data visually	
		D.5.1.1. Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)
	D.5.2. Organize visual(s) into dashboard view	

D. Tableau Visualization (Continued)	
	D.6. Polish Data / Tables
	D.6.1. Add clear titles for all visuals (size 15 font, bold)
	D.6.2. Add single-sentence explanations for visuals underneath title (size 12) where needed
	D.6.3. Add KIPP Texas graphics and fonts
E. Written Report	
	E.1. Introduction
	E.1.1. Historical background regarding college and success
	E.1.2. Information about KIPP, including College Match Strategy
	E.2. Methods
	E.2.1. Data Sources
	E.2.1.1. List Data Sources in table
	E.2.2. Data Gathering
	E.2.2.1. Complete documentation on data gathering
	E.2.3. Data Preparation
	E.2.4. Pre-processing
	E.2.5. Narrowing data features/fields
	E.3. Results – Executive Summary
	E.3.1. Summary of key findings from research questions and models, along with question-specific takeaways and insights

E. Written Report (Continued)	
	E.4. Discussion – Research Questions
	E.4.1. Q1—Are there better college retention rates or ECC for in-state / out-of-state?
	E.4.2. Q2—Are our kids “beating” the ECC rates? Is so, why?
	E.4.3. Q3—Is there a correlation between ECC and applying to 6+ schools, or 9+ schools? If applying to higher ECC schools, are they choosing them?
	E.4.4. Q4— What is correlation between college ready indicators (TSI, ACT, GPA, AP, 1st year GPA / retention rates, # of applications submitted) and college success? Is there any difference in 2-year and 4-year program success?
	E.5. Recommendations for future research topics / next steps
	E.6. Add Appendix with common
	E.7. Submit to client for final review and feedback

VI. Exit Criteria

Milestone	Exit Criteria
M1 - Data Cleaned	There will be 4 sets to be pre-processed. The PM will own this process, as he is closest to the organization and data sets. The end goal here should be the 4 cleaned data sets that meet the criteria listed in the WBS for pre-processing.
M2 - Research question analysis complete	<p>Four (4) separate R scripts will be written that form the basis of our research analysis, one for each of the four (4) research questions using four (4) different data sets.</p> <p>The PM and data analysts will have initial one-on-one check-ins about their research question and establish a direction and ideas for analysis methods. The PM will also set a half-way check-in with the data analysts to problem solve around any technical barriers, and pressure-test initial findings or results.</p> <p>For each R script, the data will be loaded at the top with packages installed at the top and will be formatted using R markdown files for easier sharing and organization of thoughts. Sections of the scripts will include data import, formatting, model exploration, and any written explanations or findings listed outside of the markdown script sections. # notation will also be used in code lines for brief explanations of steps.</p> <p>Finally, the R file and written reports will be reviewed by the PM and signed off by 11/24.</p>
M3 - Tableau Visualization of Data	<p>Four (4) dashboards within one (1) Tableau workbook will form the basis of this section. To highlight the results of each research question, relevant graphics will be displayed in an accessible, well-labelled dashboard view, with tabs separating the research questions.</p> <p>Visuals will be chosen based on the ability to highlight relevant trends and patterns found in the analysis and add tools for interaction. Such features will include slider filters, detailed tool-tip functionality, and the ability to sort or subset the given data.</p> <p>The PM will own the creation of this, with support from Gwaldys as a technical consultant. They will hold two (2) check-ins during this project window to test the functionality of the dashboards and look for ways to improve the features. Ultimately, the PM will sign off for this, and will preview with the client by 12/3 and reach client signoff by 12/11.</p>

Milestone	Exit Criteria
Written Report	<p>The final written report will be shared as a PDF with the client by 12/17 and will be the sum of all parts of this project.</p> <p>The PM will check-in with the communication lead twice during the window; once by 11/28 for all non-research sections, and by 12/3 once all research analysis has concluded. The communications lead will manage communicating around and organizing the transition of findings from R to the report via the data analysts.</p> <p>A draft will be shared with the client by 12/11, with the final product sharing and sign-off occurring no later than 12/17.</p>

VII. Milestones

ID	Milestone	Date
M1	Data cleaned	Saturday, 11/14/20
M2	Research question analysis complete	Tuesday, 11/24/20
M3	Tableau visualizations complete	Monday, 11/30/20
M5	Client review (visualizations) complete	Thursday, 12/3/20
M6	Client review (written report) complete	Friday, 12/11/20
M7	Written Report Submission Complete	Thursday, 12/17/20

VIII. Working Schedules

Task ID	Task Description	Dependency	Time(d)	Start	End
A.	Clean Data	-	2	11/13/2020	11/14/2020
A.1.	Import each data source into R	-	1	11/13/2020	11/13/2020
A.1.1.	Check for completeness and relevance of data	-	1	11/13/2020	11/13/2020
A.1.1.1.	Remove rows or columns with missing values	A.1.1.	1	11/13/2020	11/13/2020
A.1.1.2.	Remove columns/features that are not relevant to project	A.1.1.1.	1	11/13/2020	11/13/2020
A.1.2.	Check for accuracy of data	A.1.1.2.	1	11/14/2020	11/14/2020
A.1.2.1.	Remove any rows/observations that contain duplicates or errors	A.1.1.2.	1	11/14/2020	11/14/2020
A.1.3.	Format structure and set data types	A.1.2.1.	1	11/14/2020	11/14/2020
A.1.3.1.	Ensure data is imported and labeled correctly (e.g. numeric data is labeled as numeric) and convert character data to factors	A.1.3.	1	11/14/2020	11/14/2020
A.1.3.2.	Separate any data in columns that contain more than one feature and create new feature column(s) when necessary	A.1.3.1.	1	11/14/2020	11/14/2020
A.2.	Document the data cleaning process	A.1.	1	11/14/2020	11/14/2020
A.3.	Write cleaned data sets into files	A.2.	1	11/14/2020	11/14/2020
A.4.	Get approval from PM for cleaned data sets	A.3.	1	11/14/2020	11/14/2020
B.	Research Question Analysis	A.4	9	11/15/2020	11/24/20
B.1.	Research Question 1: Are there better college retention rates or ECC for in-state /out of state?	A.4	9	11/15/2020	11/24/20
B.1.1.	Research previous analysis of ECC vs college state residency	-	2	11/15/2020	11/16/2020
B.2.	Research Question 2: Are our kids “beating” the ECC rates? If so, are the factors noted?	A.4	9	11/15/2020	11/24/20

Task ID	Task Description	Dependency	Time(d)	Start	End
B.2.1.	Research any previous analysis of ECC and actual complete rate for colleges	-	2	11/15/2020	11/16/2020
B.3	Research Question 3: Is there a correlation between ECC and applying to 6+ schools, or 9+ schools?	A.4	9	11/15/2020	11/24/20
B.3.1.	Research previous analysis of number of applications submitted and college completion rates	-	2	11/15/2020	11/16/2020
B.4.	Research Question 4: What is correlation between college ready indicators (TSI, ACT, GPA, AP, 1st year GPA / retention rates, # of applications submitted) and college success? Is there any difference in 2-year and 4-year program success?	A.4	9	11/15/20	11/24/20
B.4.1.	Research previous analysis of number of applications submitted and college completion rates	-	2	11/15/20	11/16/20
D.	Tableau Visualization	-	15	11/15/20	11/30/20
D.1	Connect Data	A.4	1	11/15/20	11/15/20
D.1.1.	Import each dataset into Tableau (separate data sources in same workbook)	A.4	1	11/15/20	11/15/20
D.1.2.	Label Data sources by research question (RC1, RC2, RC3, RC4)	D.1.1.	1	11/15/20	11/15/20
D.2	RQ1 Visualization	-	5	11/15/20	11/20/20
D.2.1.	Explore data visually	A.4	3	11/15/20	11/17/20
D.2.1.1.	Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)	D.2.1.	3	11/15/20	11/17/20
D.3	RQ2 Visualization	-	5	11/15/20	11/20/20
D.3.1.	Explore data visually	A.4	3	11/15/20	11/17/20
D.3.1.1.	Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)	D.3.1.	3	11/15/20	11/17/20
B.1.2.	Import appropriate cleaned data set(s) into R and select target features	A.4	1	11/16/2020	11/16/2020

Task ID	Task Description	Dependency	Time(d)	Start	End
B.1.3.	Choose and perform the analysis method(s)	B.1.2.	3	11/16/2020	11/19/2020
B.1.3.1.	Visualize data and choose analysis method(s)	B.1.3.	2	11/16/2020	11/18/2020
B.2.2.	Import appropriate cleaned data set(s) into R and select target features	A.4	1	11/16/2020	11/16/2020
B.2.3.	Choose and perform the analysis methods	B.2.2.	3	11/16/2020	11/19/2020
B.2.3.1.	Visualize data and choose analysis method(s)	B.2.3.	2	11/16/2020	11/18/2020
B.3.2.	Import appropriate cleaned data set(s) into R and select target features	A.4	1	11/16/2020	11/16/2020
B.3.3.	Choose and perform the analysis methods	B.3.2.	3	11/16/2020	11/19/2020
B.3.3.1	Visualize data and choose analysis method(s)	B.3.3.	2	11/16/2020	11/18/2020
B.4.2.	Import appropriate cleaned data set(s) into R and select target features	A.4	1	11/16/20	11/16/20
B.4.3.	Choose and perform the analysis methods	B.4.2.	3	11/16/2020	11/19/2020
B.4.3.1	Visualize data and choose analysis method(s)	B.4.3.	2	11/16/20	11/18/20
E.	Written Report	-	29	11/16/20	12/17/20
E.1.	Introduction	-	8	11/16/20	11/24/20
E.1.1.	Historical background regarding college and success of graduates	-	4	11/16/20	11/20/20
E.2.	Methods	A.4	8	11/16/20	11/24/20
E.2.1.	Data Sources	A.4	8	11/16/20	11/24/20
E.2.1.1.	List Data Sources in table	A.4	8	11/16/20	11/24/20
E.2.2	Data Gathering	A.4	8	11/16/20	11/24/20
E.2.2.1.	Complete documentation on data gathering	-	8	11/16/20	11/24/20
E.2.3.	Data Preparation	A.4	8	11/16/20	11/24/20
E.2.4.	Pre-processing	B.1.3.3, B.2.3.3, B.3.3.3, B.4.3.3.	8	11/16/20	11/24/20

Task ID	Task Description	Dependency	Time(d)	Start	End
E.2.5.	Narrowing data features/fields	-	8	11/16/20	11/24/20
E.5.	Recommendations for future research topics / next steps	-	8	11/16/20	11/24/20
E.6.	Appendix	-	8	11/16/20	11/24/20
D.2.2.	Organize visual(s) into dashboard view	D.2.1.1.	2	11/18/20	11/20/20
D.3.2.	Organize visual(s) into dashboard view	D.3.1.1.	2	11/18/20	11/20/20
B.1.3.2.	Make connections to the college match strategy (Prepare for Match)	B.1.3.1.	1	11/19/20	11/20/20
B.1.3.3.	Document the data analysis process	B.1.3.2.	1	11/19/20	11/20/20
B.2.3.2.	Make connections to the college match strategy (Make A Strong Transition)	B.2.3.1.	1	11/19/20	11/20/20
B.2.3.3.	Document the data analysis process	B.2.3.2.	1	11/19/20	11/20/20
B.3.3.2.	Make connections to the college match strategy (Apply to all wish-list colleges)	B.3.3.1	1	11/19/20	11/20/20
B.3.3.3.	Document the data analysis process	B.3.3.2.	1	11/19/20	11/20/20
B.4.3.2	Make connections to the college match strategy (Apply to all wish-list colleges)	B.4.3.1	1	11/19/20	11/20/20
B.4.3.3	Document the data analysis process	B.4.3.2	1	11/19/20	11/20/20
D.4	RQ3 Visualization	-	7*	11/20/20	11/30/30
D.4.1.	Explore data visually	A.4	5*	11/20/20	11/28/20
D.4.1.1.	Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)	D.4.1.	5*	11/20/20	11/28/20
D.5	RQ4 Visualization	-	7*	11/20/20	11/30/30
D.5.1.	Explore data visually	A.4	5*	11/20/20	11/28/20
D.5.1.1.	Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)	D.5.1.	5*	11/20/20	11/28/20

Task ID	Task Description	Dependency	Time(d)	Start	End
B.1.4.	Write Executive Summary for research question 1	B.1.3.3.	3	11/21/20	11/24/20
B.1.4.1.	Choose 1-3 key insights gleaned from analysis	B.1.4.	3	11/21/20	11/24/20
B.2.4.	Write Executive Summary for research question 2	B.2.3.3.	3	11/21/20	11/24/20
B.2.4.1.	Choose 1-3 key insights	B.2.4.	3	11/21/20	11/24/20
B.3.4.	Write Executive Summary for research question 3	B.3.3.3.	3	11/21/20	11/24/20
B.3.4.1.	Choose 1-3 key insights	B.3.4.	3	11/21/20	11/24/20
B.4.4.	Write Executive Summary for research question 4	B.4.3.3	3	11/21/20	11/24/20
B.4.4.1.	Choose 1-3 key insights	B.4.4.	3	11/21/20	11/24/20
E.1.2.	Information about KIPP, including College Match Strategy	-	4	11/21/20	11/24/20
E.3.	Results – Executive Summary	B.1.4, B.2.4., B.3.4, B.4.4.	9	11/24/20	12/3/20
E.3.1.	Gather key findings from research questions and models, along with question-specific takeaways and insights	B.1.4, B.2.4., B.3.4, B.4.4.	9	11/24/20	12/3/20
E.4.	Discussion – Research Questions	B.1.4, B.2.4., B.3.4, B.4.4.	9	11/24/20	12/3/20
E.4.1.	Q1—Are there better college retention rates or ECC for in-state / out-of-state?	B.1.4	9	11/24/20	12/3/20
E.4.2.	Q2—Are our kids “beating” the ECC rates? Is so, why?	B.2.4	9	11/24/20	12/3/20
E.4.3.	Q3—Is there a correlation between ECC and applying to 6+ schools, or 9+ schools? If applying to higher ECC schools, are they choosing them?	B.3.4	9	11/24/20	12/3/20

Task ID	Task Description	Dependency	Time(d)	Start	End
E.4.4.	Q4— What is correlation between college ready indicators (TSI, ACT, GPA, AP, 1st year GPA / retention rates, # of applications submitted) and college success? Is there any difference in 2-year and 4-year program success?	B.4.4	9	11/24/20	12/3/20
E.4.5.2.	Choose 2-3 screenshots of Tableau Visualizations that capture key points or connections to research questions	D.6.3	9	11/24/20	12/3/20
D.4.2.	Organize visual(s) into dashboard view	D.4.1.1.	2	11/29/20	11/30/30
D.5.2.	Organize visual(s) into dashboard view	D.5.1.1.	2	11/29/20	11/30/30
D.6.	Polish Data / Tables	-	1	11/29/20	11/29/20
D.6.1.	Add clear titles for all visuals (size 15 font, bold)	D.2.2, D.3.2, D.4.2, D.5.2	1	11/29/20	11/29/20
D.6.2.	Add single-sentence explanations for visuals underneath title (size 12) where needed	D.6.1.	1	11/29/20	11/29/20
D.6.3.	Add KIPP Texas graphics and fonts	D.6.2.	1	11/29/20	11/29/20
E.7.	Submit to client for final review and feedback	E.6	1	12/11/20	12/11/20
E.8.	Implement Client Feedback and share final product	E.7.	6	12/11/20	12/17/20

*No scheduled work between 11/25 - 11/27 for Fall (Thanksgiving) Break

IX. Required Resources

PEOPLE		
Name	Email	Project Role
Carmella Blackwell	blkwlldsgns@gmail.com	Communications Lead
Gwladys Peg	gwladys.peg1@gmail.com	Data Analyst
Melvin Zaldivar	mzaldiva2@gmail.com	Data Analyst
Taylor Gilfillan	gilfillant@gmail.com	Project Manager

TECHNOLOGY	
Equipment / Tool	Description
R Studio	RStudio is an integrated development environment for R, a programming language for statistical computing and graphics that will be used in the research analysis aspect of the project.
Tableau	Tableau is the data visualization software that will be used to create the workbook that will be shared via SFTP to KIPP Texas.
Salesforce	Student and alum data are warehoused in Salesforce, which is an online database management system solution. Student and alum data are accessed here, and access and storage meet the guidelines set forth by the signed Data Usage Agreement (DUA) between KIPP Texas and UHD.
Google Suite	A cloud-based-workplace (Google product) used in the project to plan, communicate, create documents, store data and any other pertinent information for the final project.
Monday.com	Customizable workflow space used to track tasks and assignments to complete the final project in a timely manner.

X. Major Contributors

P = Primary Responsibility | S = Support | Blank = No Responsibility

Task ID	Task Description	Taylor	Carmella	Melvin	Gwladys
A.	Clean Data	P			
A.1.	Import each data source into R	P			
A.1.1.	Check for completeness and relevance of data	P			
A.1.1.1.	Remove rows or columns with missing values	P			
A.1.1.2.	Remove columns/features that are not relevant to project	P			
A.1.2.	Check for accuracy of data	P			
A.1.2.1.	Remove any rows/observations that contain duplicates or errors	P			
A.1.3.	Format structure and set data types	P			
A.1.3.1.	Ensure data is imported and labeled correctly (e.g. numeric data is labeled as numeric) and convert character data to factors	P			
A.1.3.2.	Separate any data in columns that contain more than one feature and create new feature column(s) when necessary	P			
A.2.	Document the data cleaning process	P			
A.3.	Write cleaned data sets into files	P			
A.4.	Get approval from PM for cleaned data sets	P			
B.	Research Question Analysis			P	
B.1.	Research Question 1: Are there better college retention rates or ECC for in-state /out of state?			P	
B.1.1.	Research previous analysis of ECC vs college state residency			P	
B.1.2.	Import appropriate cleaned data set(s) into R and select target features			P	
B.1.3.	Choose and perform the analysis method(s)			P	S
B.1.3.1.	Visualize data and choose analysis method(s)			P	S
B.1.3.2.	Make connections to the college match strategy (Prepare for Match)			P	S
B.1.3.3.	Document the data analysis process			P	S

Task ID	Task Description	Taylor	Carmella	Melvin	Gwladys
B.1.4.	Write Executive Summary for research question 1			P	S
B.1.4.1.	Choose 1-3 key insights gleaned from analysis			P	S
B.2.	Research Question 2: Are our kids “beating” the ECC rates? If so, are the factors noted?			P	
B.2.1.	Research any previous analysis of ECC and actual complete rate for colleges			P	
B.2.2.	Import appropriate cleaned data set(s) into R and select target features			P	
B.2.3.	Choose and perform the analysis methods		S	P	
B.2.3.1.	Visualize data and choose analysis method(s)		S	P	
B.2.3.2.	Make connections to the college match strategy (Make A Strong Transition)		S	P	
B.2.3.3.	Document the data analysis process		S	P	
B.2.4.	Write Executive Summary for research question 2		S	P	
B.2.4.1.	Choose 1-3 key insights		S	P	
B.3	Research Question 3: Is there a correlation between ECC and applying to 6+ schools, or 9+ schools?				P
B.3.1.	Research previous analysis of number of applications submitted and college completion rates				P
B.3.2.	Import appropriate cleaned data set(s) into R and select target features				P
B.3.3.	Choose and perform the analysis methods			S	P
B.3.3.1	Visualize data and choose analysis method(s)			S	P
B.3.3.2.	Make connections to the college match strategy (Apply to all wish-list colleges)			S	P
B.3.3.3.	Document the data analysis process			S	P
B.3.4.	Write Executive Summary for research question 3			S	P
B.3.4.1.	Choose 1-3 key insights			S	P
B.4.	Research Question 4: What is correlation between college ready indicators (TSI, ACT, GPA, AP, 1st year GPA / retention rates, # of applications submitted) and college success? Is there any difference in 2-year and 4-year program success?		P		

Task ID	Task Description	Taylor	Carmella	Melvin	Gwladys
B.4.1.	Research previous analysis of number of applications submitted and college completion rates		P		
B.4.2.	Import appropriate cleaned data set(s) into R and select target features		P		
B.4.3.	Choose and perform the analysis methods	S	P		
B.4.3.1	Visualize data and choose analysis method(s)	S	P		
B.4.3.2	Make connections to the college match strategy (Apply to all wish-list colleges)	S	P		
B.4.3.3	Document the data analysis process	S	P		
B.4.4.	Write Executive Summary for research question 4	S	P		
B.4.4.1.	Choose 1-3 key insights	S	P		
D.	Tableau Visualization	P			
D.1	Connect Data	P			
D.1.1.	Import each dataset into Tableau (separate data sources in same workbook)	P			
D.1.2.	Label Data sources by research question (RC1, RC2, RC3, RC4)	P			
D.2	RQ1 Visualization	P			
D.2.1.	Explore data visually	P			S
D.2.1.1.	Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)	P			S
D.2.2.	Organize visual(s) into dashboard view	P			S
D.3	RQ2 Visualization	P			
D.3.1.	Explore data visually	P			S
D.3.1.1.	Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)	P			S
D.3.2.	Organize visual(s) into dashboard view	P			S
D.4	RQ3 Visualization	P			
D.4.1.	Explore data visually	P			
D.4.1.1.	Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)	P			

Task ID	Task Description	Taylor	Carmella	Melvin	Gwladys
D.4.2.	Organize visual(s) into dashboard view	P			
D.5	RQ4 Visualization	P			
D.5.1.	Explore data visually	P			
D.5.1.1.	Decide on visual type(s) to capture trend (scatter plot, bar chart, geomapping, table, time series, histogram, treemaps)	P			
D.5.2.	Organize visual(s) into dashboard view	P			
D.6.	Polish Data / Tables	P			
D.6.1.	Add clear titles for all visuals (size 15 font, bold)	P			S
D.6.2.	Add single-sentence explanations for visuals underneath title (size 12) where needed	P			S
D.6.3.	Add KIPP Texas graphics and fonts	P			S
E.	Written Report	S	P		S
E.1.	Introduction		P		S
E.1.1.	Historical background regarding college and success of graduates		P		S
E.1.2.	Information about KIPP, including College Match Strategy		P		S
E.2.	Methods		P		S
E.2.1.	Data Sources		P		S
E.2.1.1.	List Data Sources in table		P		S
E.2.2.	Data Gathering		P		S
E.2.2.1.	Complete documentation on data gathering		P		S
E.2.3.	Data Preparation		P		S
E.2.4.	Pre-processing		P		S
E.2.5.	Narrowing data features/fields		P		S
E.3.	Results – Executive Summary	S	P		
E.3.1.	Gather key findings from research questions and models, along with question-specific takeaways and insights	S	P		
E.4.	Discussion – Research Questions	S	P		
E.4.1.	Q1—Are there better college retention rates or ECC for in-state / out-of-state?	S	P		
E.4.2.	Q2—Are our kids “beating” the ECC rates? Is so, why?	S	P		

Task ID	Task Description	Taylor	Carmella	Melvin	Gwladys
E.4.3.	Q3—Is there a correlation between ECC and applying to 6+ schools, or 9+ schools? If applying to higher ECC schools, are they choosing them?	S	P		
E.4.4.	Q4— What is correlation between college ready indicators (TSI, ACT, GPA, AP, 1st year GPA / retention rates, # of applications submitted) and college success? Is there any difference in 2-year and 4-year program success?	S	P		
E.4.5.2.	Choose 2-3 screenshots of Tableau Visualizations that capture key points or connections to research questions	S	P		
E.5.	Recommendations for future research topics / next steps		P		S
E.6.	Appendix	S	P		
E.7.	Submit to client for final review and feedback	S	P		S
E.8	Implement Client Feedback and share final product	S	P		S

***Task group C was removed from the project scope due to updates from client**

XI. Control Systems

To minimize the gap between our project planning and project execution in order to achieve our project objective, our control system will focus on our schedules and deliverables.

1. Requirements:

- a. **Inputs:** Ensure data accuracy, and client's expectations are feasible and realistic.
- b. **Tools and techniques:**
 - i. Use a project management software for scheduling.
 - ii. Use a statistical computing tool for data analysis.
 - iii. Use a data visualization tool to translate queries into visualization.
 - iv. Ways of working: Set schedules / timeframe, inspection, meetings with team, client, and experts.

2. **Outputs:** Ensure product (deliverables) integrity, produce a final report with an on-schedule delivery to the client.

Quality Control

Our project quality control is the steps taken to monitor our results to see if they conform to requirements.

1) Quality Control Process:

- a) **Work performance Report:** Each member of the team meets one-on-one weekly with the Project Manager on Monday evenings. The project manager provides recommendations and updates to the team during scheduled meetings or whenever need be. Each team member makes observations, justifications, and provides information notes to the project manager and any other member of the team.
- b) The project manager uses monday.com dashboard to follow up the status of the work of each team member.
- c) Every analyst uses a dashboard / checklist to ensure the ask is fully met before delivery.

2) Iterative control (cross checking / examination control):

- a) Team members cross check each other to ensure we all have the same understanding of the work being done. Data Analysts cross-check each other's code for debug errors or statistical errors halfway through analysis.

- b) Project manager to review and cross check the quality of data before putting it into production by having a minimum of 2 check-ins per task group during development to course correct towards exit criteria for upcoming milestones.

3) Change request:

- a) This is the documentation used to request a change that includes the request, the reason for the request, the condition of success, the expected completion date, and the expected value.
- b) Data analysts submit a request form for either a corrective action (it is an intentional activity that's realigns the performance of the project work), a preventive action (it is an intentional activity that ensure a future performance of the project work), or a defective repair (it is an intentional activity to modify a non-conforming product).
- c) The project manager reviews the request, discusses the impact with the client who will then formerly provide his decision.

4) Expert Judgment:

- a) Project manager ensures the quality of data to work with by getting it directly from a qualified designated resource by the client.
- b) She/He reviews the final product before delivery. She/He ensures our deliverables are in line with the initial given outline by presenting on a biweekly basis our work status to our project management professor and the final project as well. She/He also previews the work with the client to make any adjustments before delivery.

Change Control

Our project change control is based on the six-steps process for change control.

1) Change control process:

- a. Project Manager enters initial change control information into our change control log.
- b. Project Manager determines if the change should be processed.
- c. Project Manager submits recommendations to our client for review and approval.
- d. Project Manager updates the project plan.
- e. Project Manager distributes the updated plan.
- f. Project Manager monitors the change and tracks progress against the revised plan.

Conflict Control

Although communication is key, one way to make conflict management easier is for the project manager as well as every team member to invest themselves in truly understanding our project constituents. Moreover, Each individual in the project including the project manager and the client should be able to listen effectively, identify specific points of disagreements, clearly express his / her needs, be open to feedback, view conflict as an opportunity for growth, and treat each other respectfully.

Our conflict resolution process will be based on collaborating, compromising, and accommodating.

Conflict resolution process:

- 1) **Collaborating:** The project manager effectively communicates with each party involved to understand their perspective and gain insights that will help him find a win-win solution that satisfies all parties involved.
- 2) **Compromising:** The project manager mediates between the conflict parties to find a mutually acceptable resolution that partly satisfies the concerns of each party involved.
- 3) **Accommodating:** If the project manager determines there is a minor issue with the client that will not have lasting consequences, he will accommodate the client by stretching his plans a little. The same process will apply to team members for example in the case some is sick or has personal issues, the rest of the team will accommodate him.

XII. Risk Areas with Contingencies

RISK	PROBABILITY	IMPACT	TRIGGER	MITIGATING ACTION(S)
Struggle to connect to practical solutions (How do we translate results to someone without the data analytics background)	Medium	High	Meeting with client a week prior to final presentation and client not quite understanding the tableau workbook	Meet with PM and lean in with client for support with to translate model/results to understandable insight
Parts of the results from the data come back inconclusive and it hinders on the written report	Low	High	R Model/Indicators show no correlation or relationship to any of the data attributes	Team will meet with PM to ask for suggestions with regards to the data and ask for additional resources to be able to improve the model
Running into technical difficulties with software program(s)	Medium	High	Tableau Workbook is run 1-2 days prior to the project due date, and it doesn't run properly	Member(s) in charge of the Tableau workbook will notify the PM of problems, and PM and other member(s) will troubleshoot the software to identify the issue before the project due date.
A member (family/relative) falls ill or has personal issues that requires them to be out for several weeks	Medium	Medium	Task assigned to group member(s) is not completed due to illness/personal issues 1-2 days of due date.	PM will support and reassign other member(s) to assist in the task