

1) Introduction

I created a database of US Presidencies. My main tables are: US Presidents, Vice Presidents, Presidencies, Vice Presidencies, and Universities. I created three linking tables to handle many-to-many relationships. I thought this topic would be relatively straightforward, but the complexities of the US presidential system had to be accounted for in my databases.

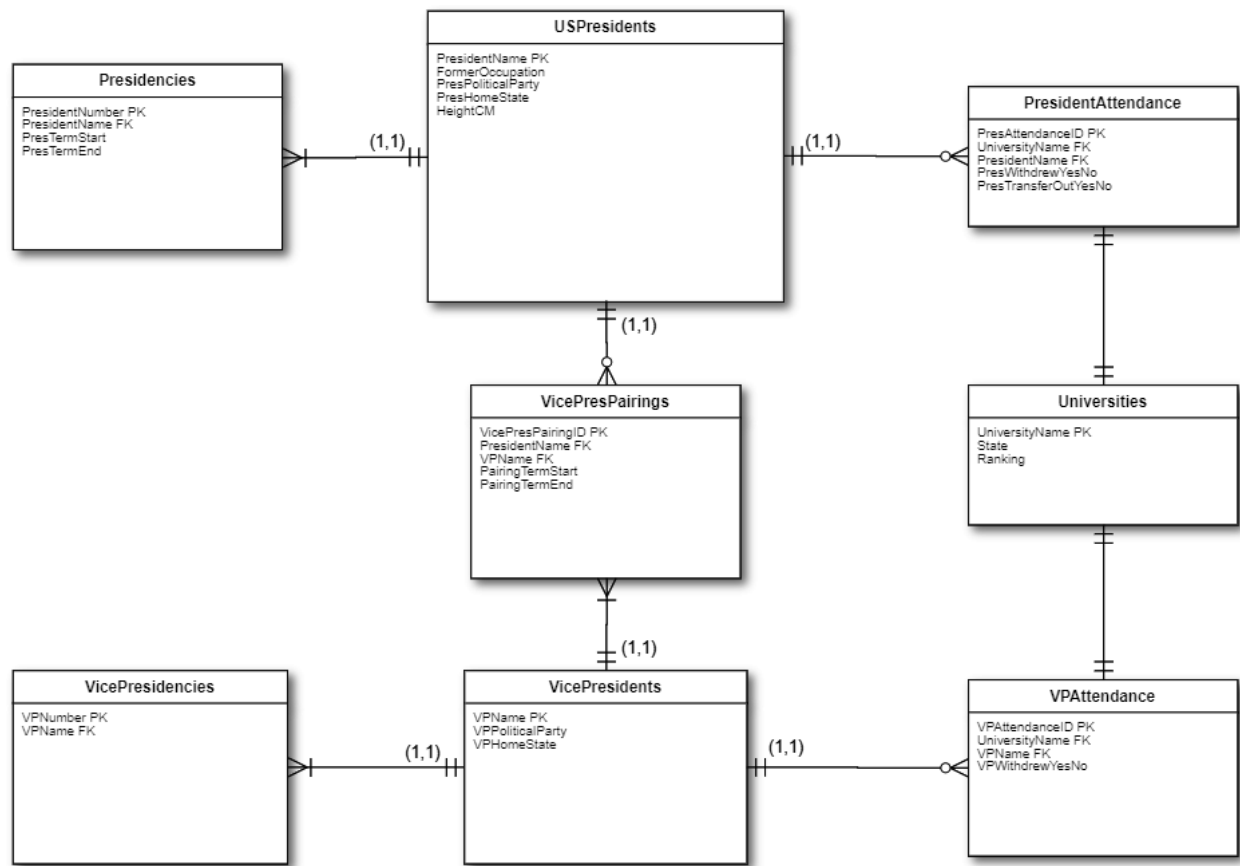
Initially I attempted to build a database of video game sales, but it was far too complex for me at this point. The topic of US presidents ended up being a better way for me to work through common mistakes, errors, and bugs. Also, the topic is probably familiar to everyone, so it should be easier to present my results.

2) Mission Statement and Mission Objectives:

- Provide researchers quick access data about US presidents and vice presidents.
- Create insightful views that demonstrate a presence or absence of trends/correlations or reveal some interesting quirks in the institution of the US presidency.
- Information would only need to be updated when the current president or vice president changes.
- A complete version of this project would have an online interactive database where users can look up specific information or build their own custom views.
- The database would be useful for journalists, historians, political scientists, and anyone interested in US history.

3) Conceptual Design:

- E-R diagram, with the final table names and fields
- Identifying primary and foreign keys
- Create links between tables
- Label the type of relationship and degree of participation



4) Logical design:

- Field specifics for all tables, including linking tables
- Identifying primary keys and foreign keys if any
- For each field, list the table that the field belongs to, field name, data type, data length, null determination, index, key type, default value, range of value and description [All

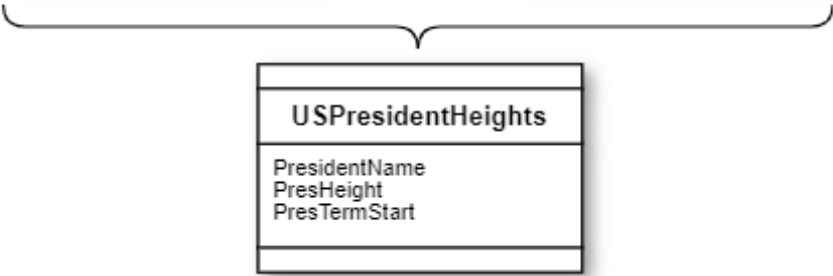
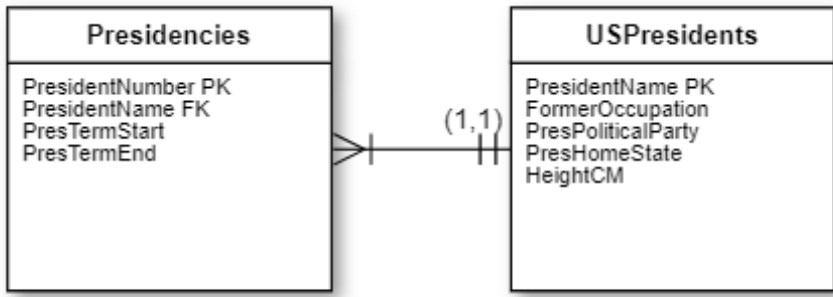
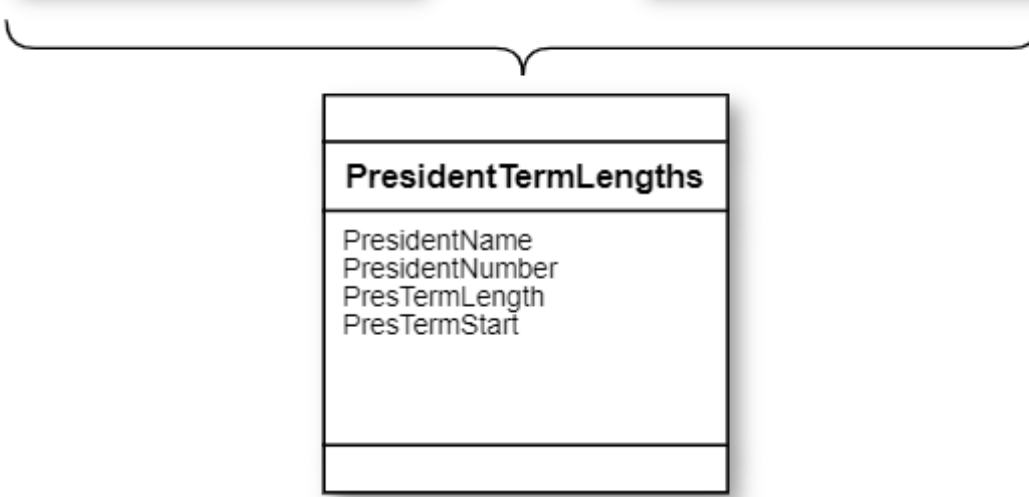
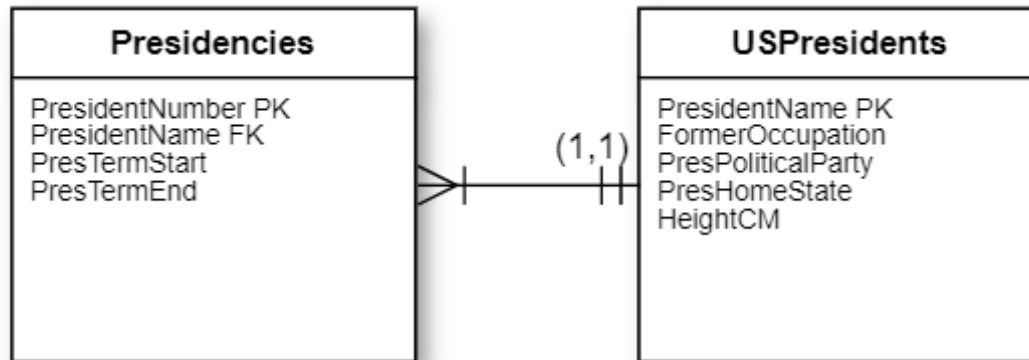
data documentation items]

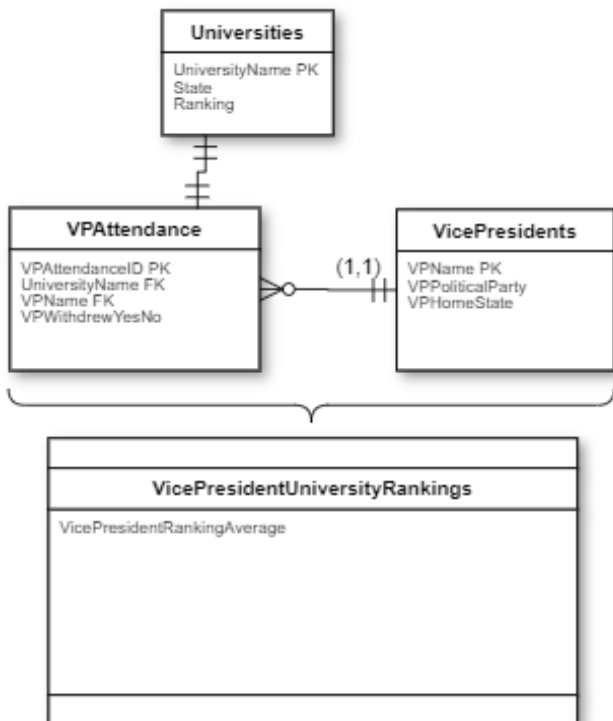
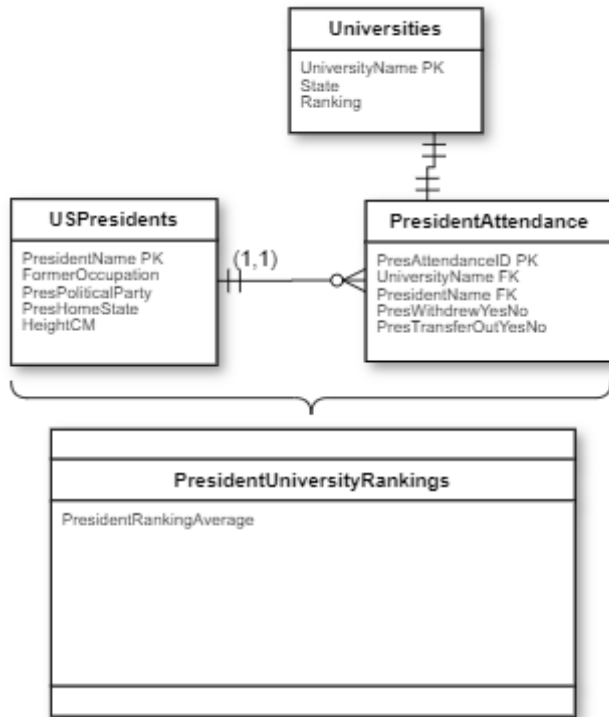
	Field Name	Description	Key Type	Data Type	Length	Null Support	Index	Default Value	Value Range
USPresidents	PresidentName	The president's full name.	PK	VarChar	25	NOT NULL	Yes		
	FormerOccupation	The president's job before becoming president.		Text		NULL			
	PresPoliticalParty	The president's political party.		VarChar	25	NULL	Yes		
	PresHomeState	The president's home state.		VarChar	50	NOT NULL	Yes		
	HeightCM	The president's height in cm.		Int		NOT NULL	Yes		
Presidencies	PresidentNumber	The president's numerical order in history.	PK	Int		NOT NULL	Yes		
	PresidentName	The president's full name.	FK	VarChar	25	NOT NULL			
	PresTermStart	The starting year of the president's term(s).		Int		NOT NULL	Yes		
	PresTermEnd	The end of the president's term(s).		Int		NULL	Yes		
VicePresidents	VPName	The vice president's full name.	PK	VarChar	25	NOT NULL	Yes		
	VPPoliticalParty	The vice president's political party.		VarChar	25	NULL	Yes		
	VPHomeState	The vice president's home state.		VarChar	50	NOT NULL	Yes		
VicePresidencies	VPNumber	The vice president's numerical order in history.	PK	Int		NOT NULL	Yes		
	VPName	The vice president's full name.	FK	VarChar	25	NOT NULL			
VicePresPairings	VicePresPairingID	An arbitrary pairing ID for the president and the vice president.	PK	Int		NOT NULL	Yes		
	PresidentName	The president's full name.	FK	VarChar	25	NOT NULL			
	VPName	The vice president's full name.	FK	VarChar	25	NOT NULL			
	PairingTermStart	The starting year of the pairing's term(s) together.		Int		NOT NULL	Yes		
	PairingTermEnd	The ending year of the pairing's term(s) together.		Int		NULL	Yes		
Universities	UniversityName	The name of the university.	PK	VarChar	200	NOT NULL	Yes		
	State	The state where the university's campus is located.		VarChar	50	NOT NULL			
	Ranking	The university's Forbes 2019 ranking.		Int		NOT NULL	Yes		
PresidentAttendance	PresAttendanceID	An arbitrary ID for the president's attendance at a university.	PK	Int		NOT NULL	Yes		
	UniversityName	The name of the university that the president attended.	FK	VarChar	200	NOT NULL			
	PresidentName	The president's full name.	FK	VarChar	25	NOT NULL			
	PresWithdrawYesNo	Whether or not the president withdrew from the university.		VarChar	3	NOT NULL	Yes		"Yes" or "No"
	PresTransferOutYesNo	Whether or not the president transferred out of the university.		VarChar	3	NOT NULL	Yes		"Yes" or "No"
VPAffendance	VPAffendanceID	An arbitrary ID for the vice president's attendance at a university.	PK	Int			Yes		
	UniversityName	The name of the university that the vice president attended.	FK	VarChar	200	NOT NULL			
	VPName	The vice president's full name.	FK	VarChar	25	NOT NULL			
	VPWithdrawYesNo	Whether or not the vice president withdrew from the university.		VarChar	3	NOT NULL	Yes		"Yes" or "No"

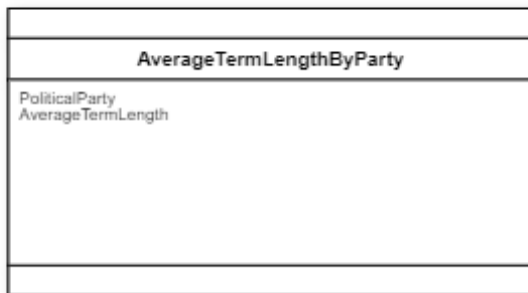
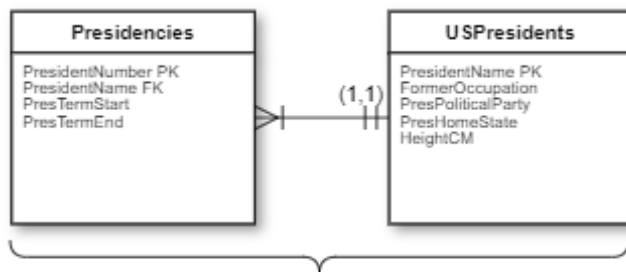
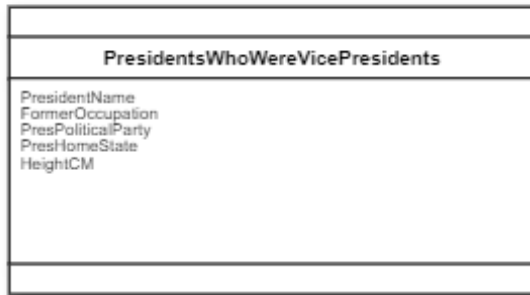
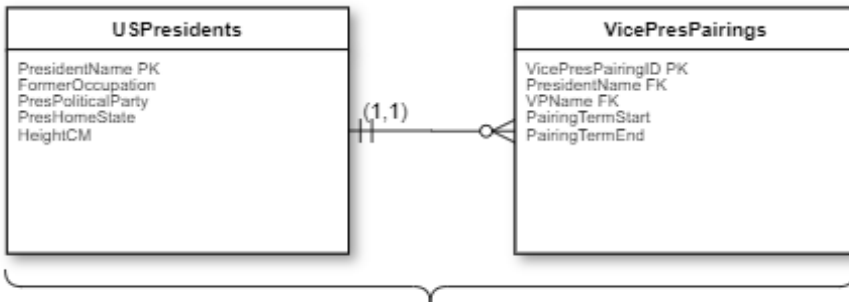
· Specify business rules, if any

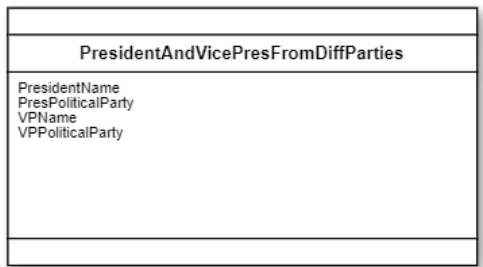
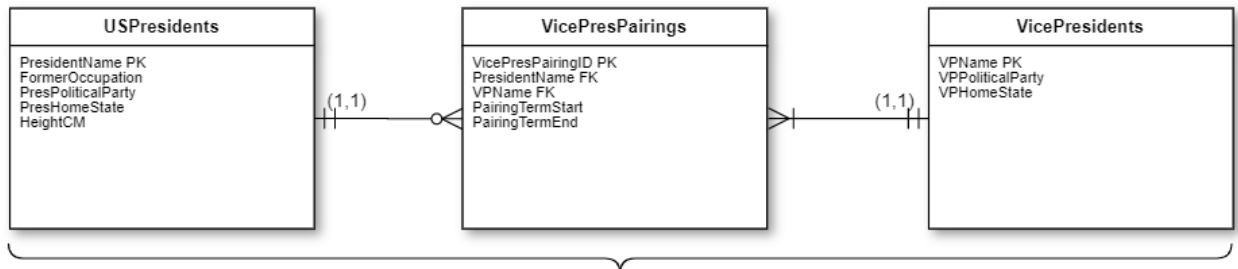
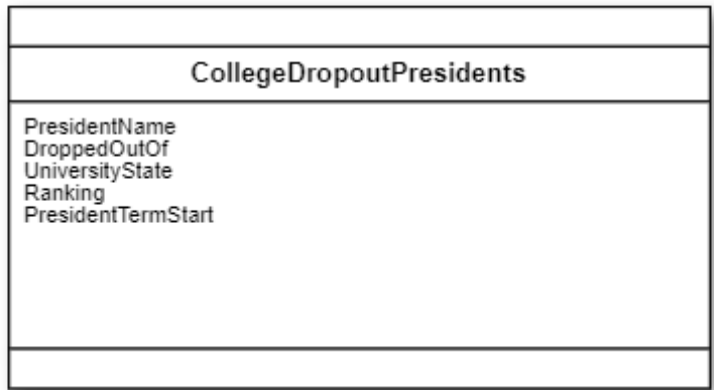
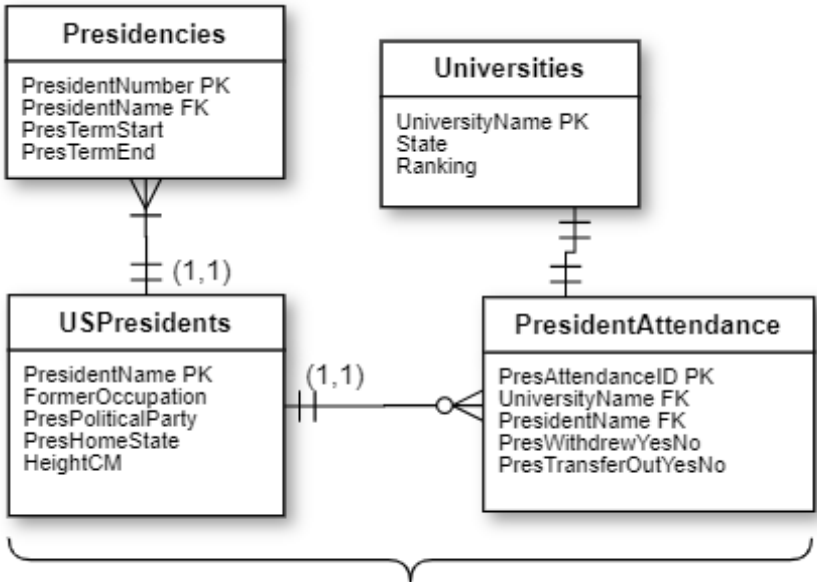
- Every vice president must have at least once president.
- Presidents don't need to have a vice president.
- Vice presidents can later become presidents.
- Presidents can later become vice presidents.
- Neither presidents nor vice presidents need to have attended a university.
- Presidents and vice presidents can have attended multiple universities.
- Presidents can be more than one president number if they serve non-consecutive terms, such as Grover Cleveland being the 22nd and 24th president.
- Presidents and vice presidents must have a term start, but term end can be null in the case of terms being currently served.

· Define views, if any









Interesting problems I encountered:

- Year() won't allow year dates before 1901, so I had to have my term start and end years as Int instead.
- Repeatedly had issues with the president full names as the primary keys due to variant uses of the middle initial or full middle name.
 - If I were to start this project over, I would separate the first and last names, and use a meaningless unique ID number as the primary keys.
 - Middle names are necessary for many cases, such as George W. Bush vs George H.W. Bush, so I would also include a MiddleName column.

Lessons Learned:

- Better to overestimate the field sizes by a factor of 3 or so.
- Should just copy-and-paste lists from online sources such as Wikipedia, as that can insert nonstandard characters such as longer dashes or spaces.
- Planning things out with the ER diagrams is extremely helpful.
- Must account for how data could be entered in the future (e.g. a vice president may have non-consecutive terms in the future, so need a linking table between vice presidents and vice presidencies).

Data/Features to Add for Improvement:

- Couldn't easily find vice president heights, but that would be an interesting comparison.
- Familial relationships between presidents (father-son, siblings, cousins, etc.) are publicly available, but seemed rather complicated to put into a database.
- Adding fields for assassinations, involvement in wars, or economic trends for each presidency would be interesting.

Sources:

US University Rankings From Forbes' THE LIST: AMERICA'S TOP COLLEGES 2019:

- <https://www.forbes.com/top-colleges/#fa4c47d19877>

US Presidents List from Kaggle user HARSHITA GUPTA:

- <https://www.kaggle.com/harshitagpt/us-presidents>

US President Heights from Kaggle user Zurda:

- <https://www.kaggle.com/hakabuk/us-presidents-heights-how-low-can-u-go>

Wikipedia President and Vice President Lists:

- https://en.wikipedia.org/wiki/List_of_presidents_of_the_United_States_by_education
- https://en.wikipedia.org/wiki/List_of_vice_presidents_of_the_United_States_by_education
- https://en.wikipedia.org/wiki/List_of_vice_presidents_of_the_United_States_by_place_of_primary_affiliation