

Tennis Analysis

Ariana Zuberovic

BDA 761

- Historical data from the ATP
 - “Association of Tennis Professionals”
 - Network of men’s tennis players formed by tennis players
- Years 1968-2015
- Attributes include:
 - Tournament Name
 - Ex: US Open
 - Surface
 - Ex: Grass
 - Score
 - Ex: 6-3 6-2 6-4
 - Player name, country and ID number

Data

- Historical Data, Open Data
- Available on GitHub
- JeffSackmann/tennis_atp
- Yearly .csv file
- Data Prior to 1983:
 - Missing Ranks (intermittent)
 - Missing Columns:
 - Number of Aces
 - Minutes Played

Data Source

- Why has the popularity of tennis in America been facing a downward slope?
- Suspected reasons:
 - 1968 - "Open Era" Prior to this tennis players could not make a living through tennis.
 - "Rich-man's country club sport"
 - 1968- is that the US Open was broadcast on television for the first time (by CBS)

Why Tennis Data?

- Analytics of the data by decade
 - Winning Countries?
 - Nationalism has been linked to popularity among tennis, where has America fallen?
 - Attributes of the top countries?
 - Ranks
 - Grand Slams
 - Surface
 - Which surfaces do the top countries dominate?
 - Age
 - When is a players peak performance?
- Making of a top country for the next decade.

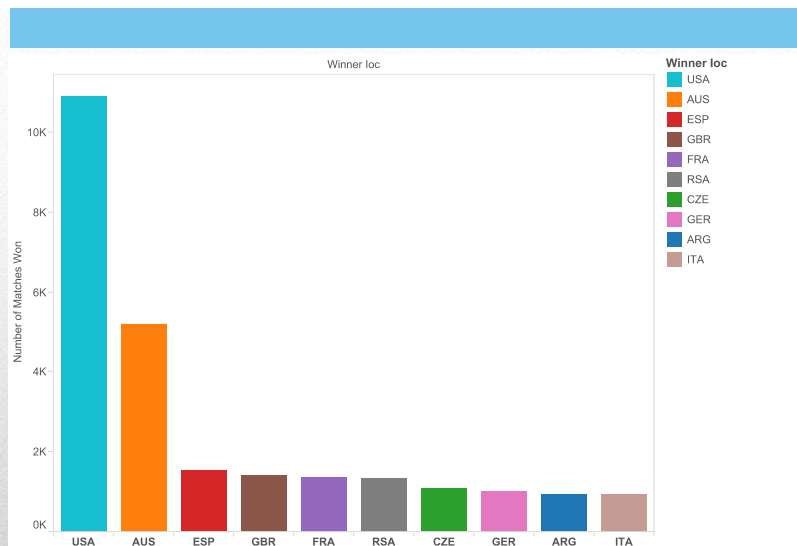
What about it?

- Missing Data
- Dates of the tournaments
 - Inaccuracy: Each match date is recorded on the same day, of course this is not possible as matches are played regularly for a series of weeks
- Columns are not completely identical
 - Can be an issue when combining data sets

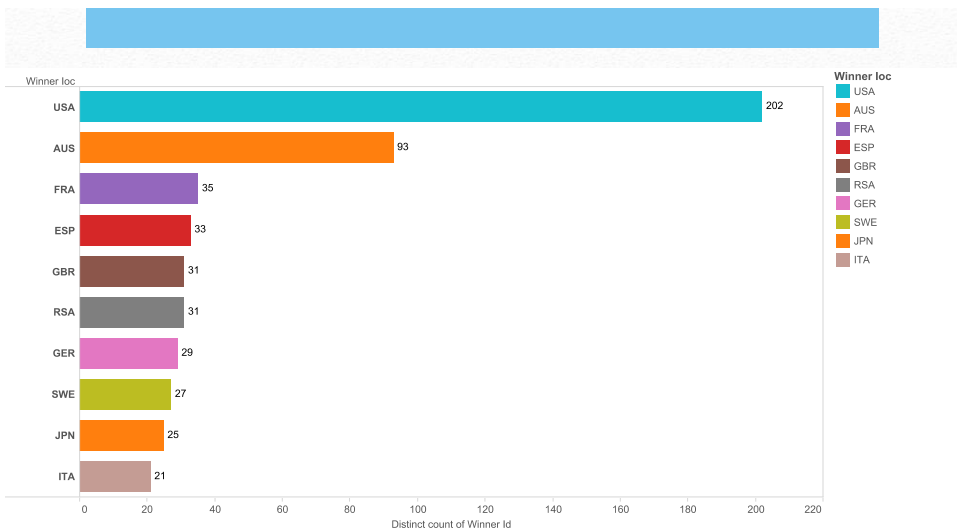
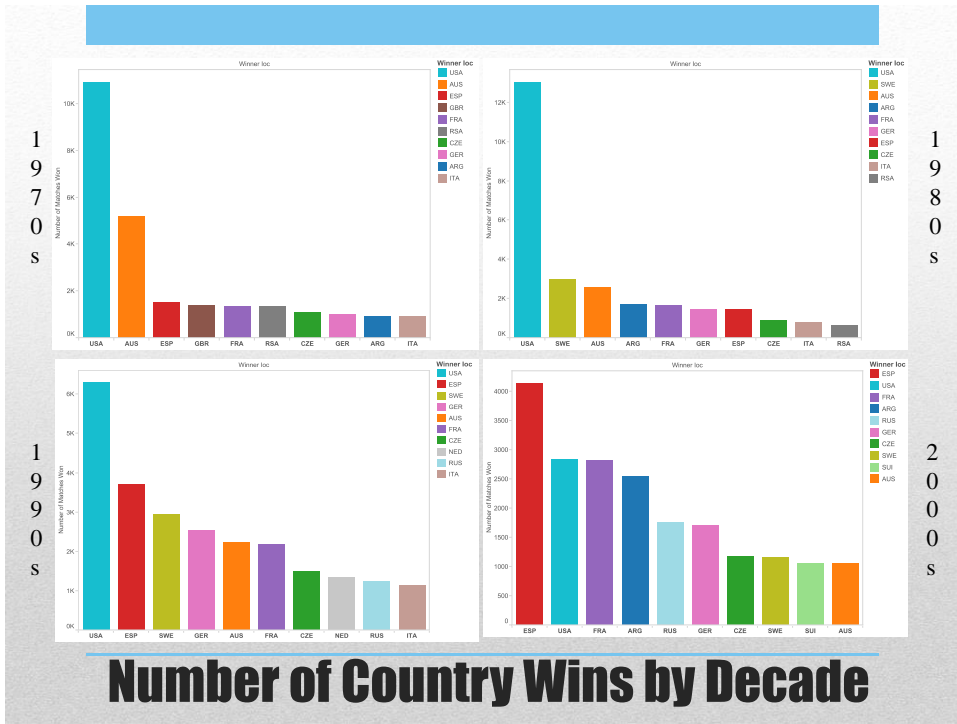
Inaccuracies with the Data Sets

- Decade Data
 - 1970s
 - 1980s
 - 1990s
 - 2000s
- Tableau
 - Extract and append data from separate files
 - Ensure: Data Columns are **Identical**
- Calculated fields for
 - Weighted number of Winning Countries
 - Restricted Ranking Numbers

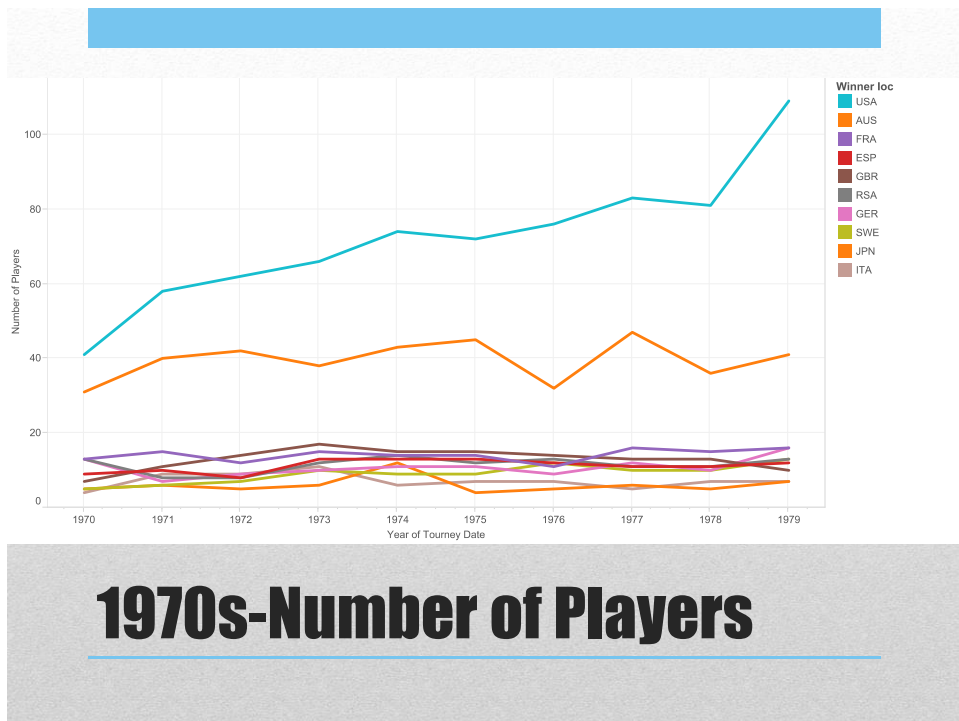
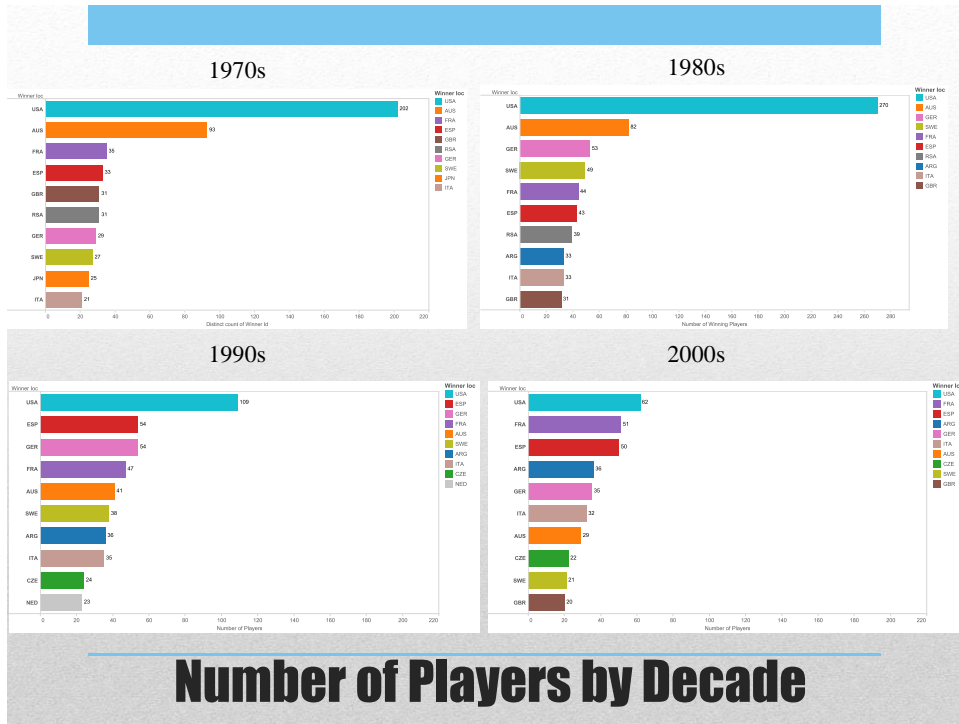
Organization

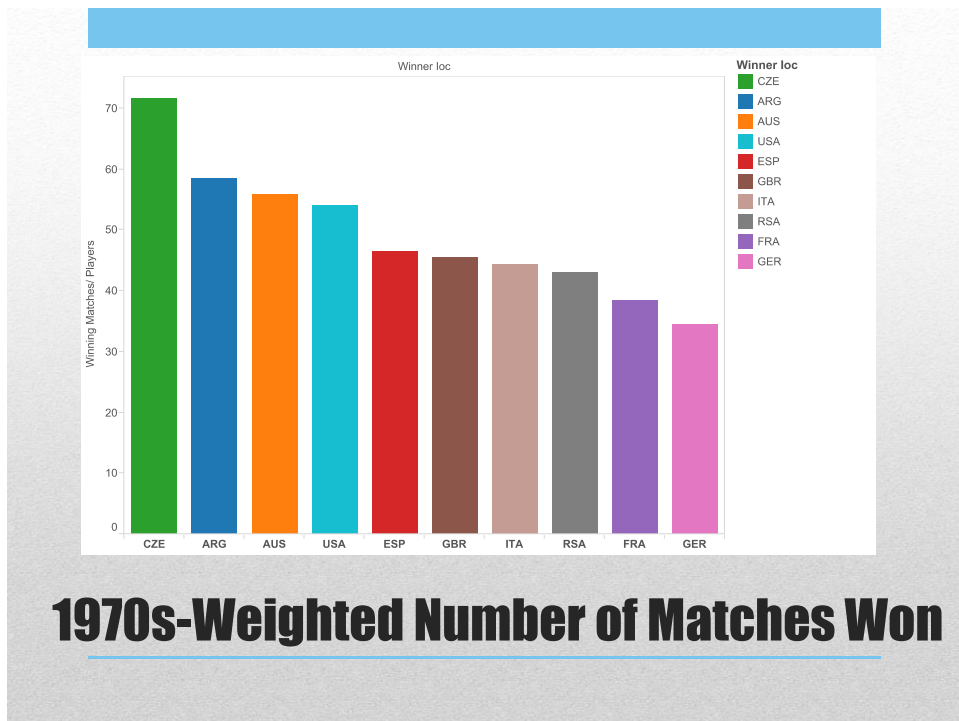
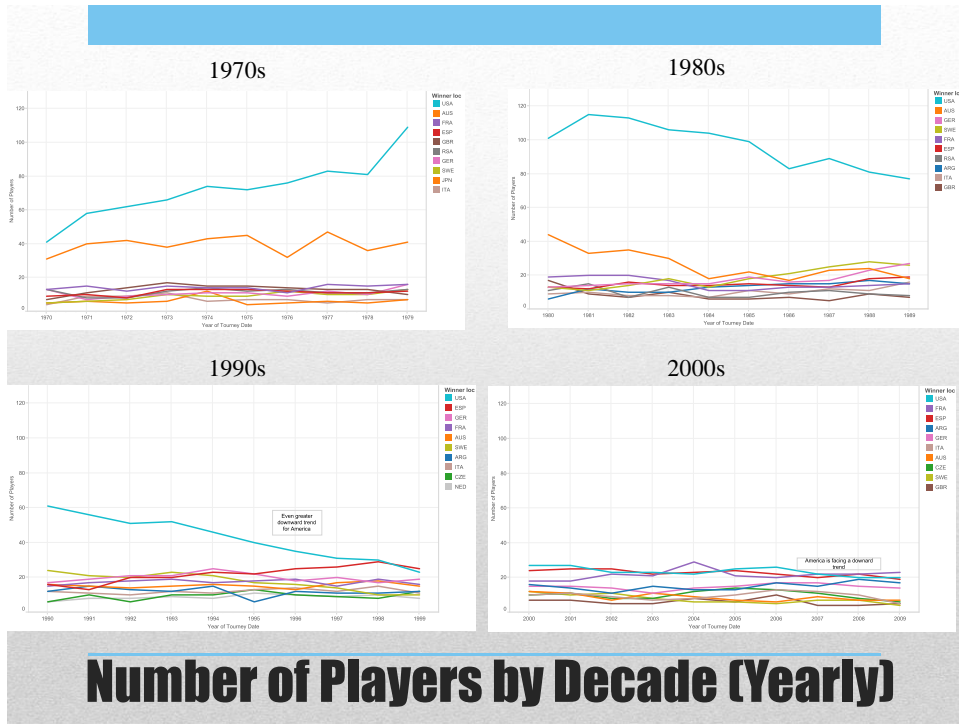


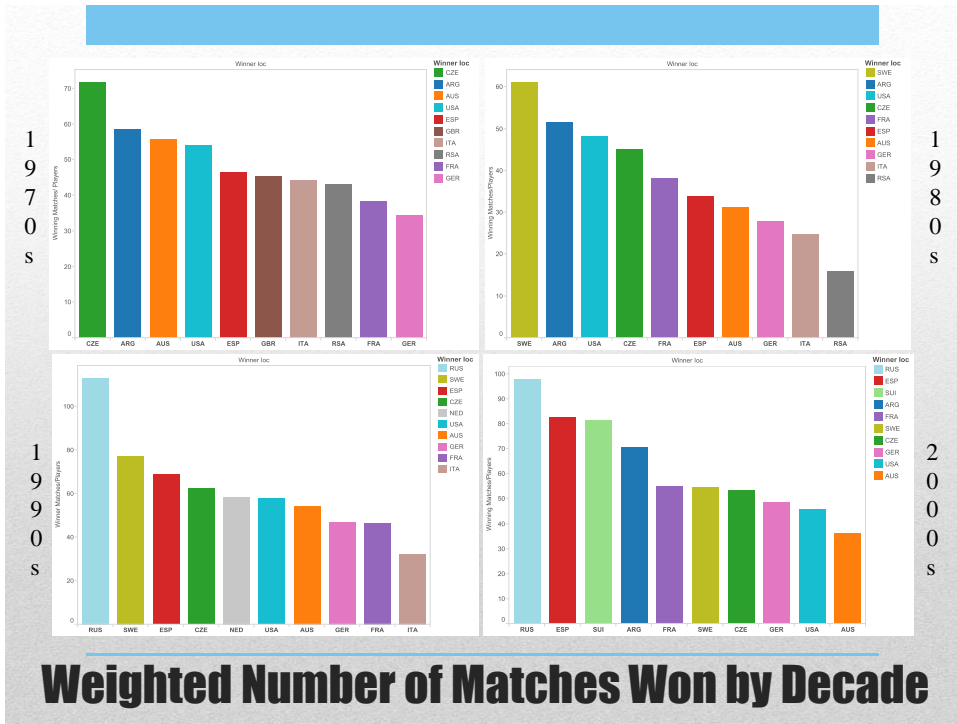
1970s-Matches Won



1970s-Number of Players

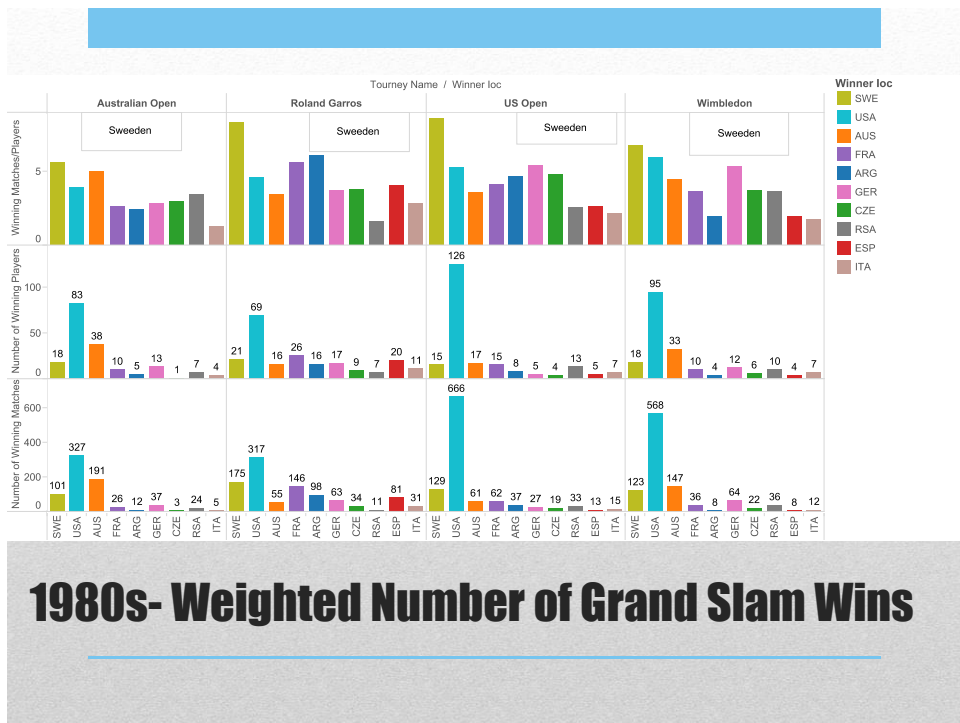
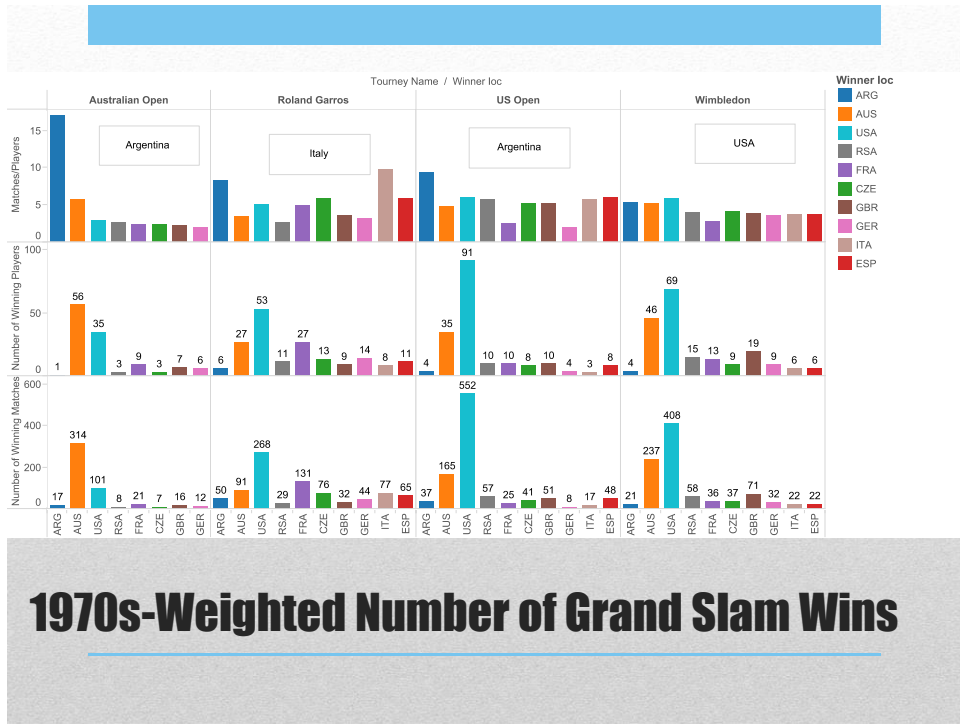


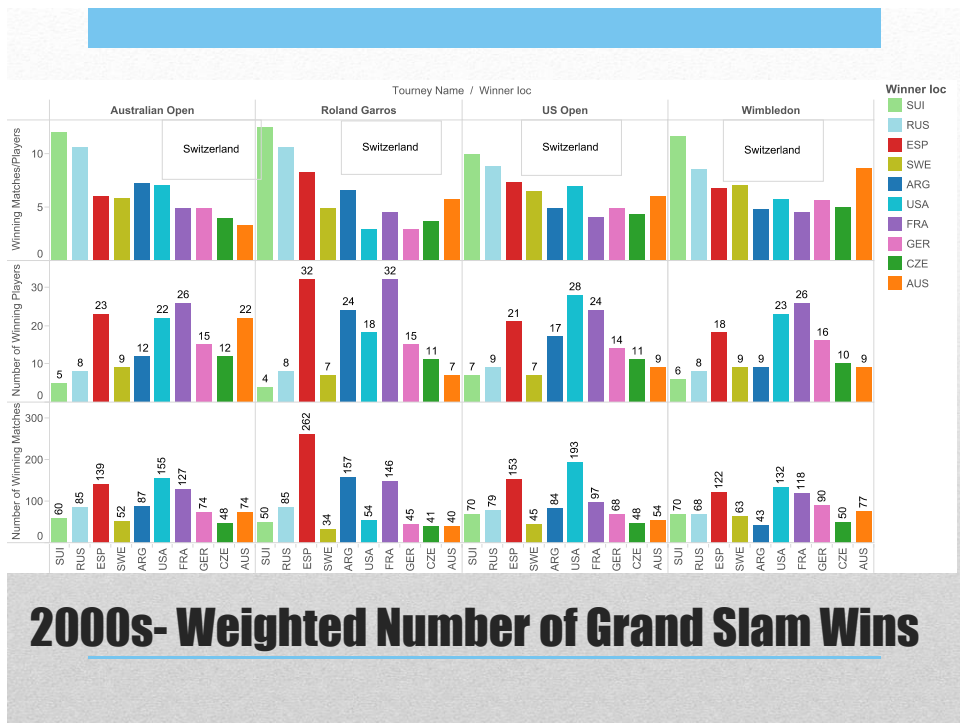
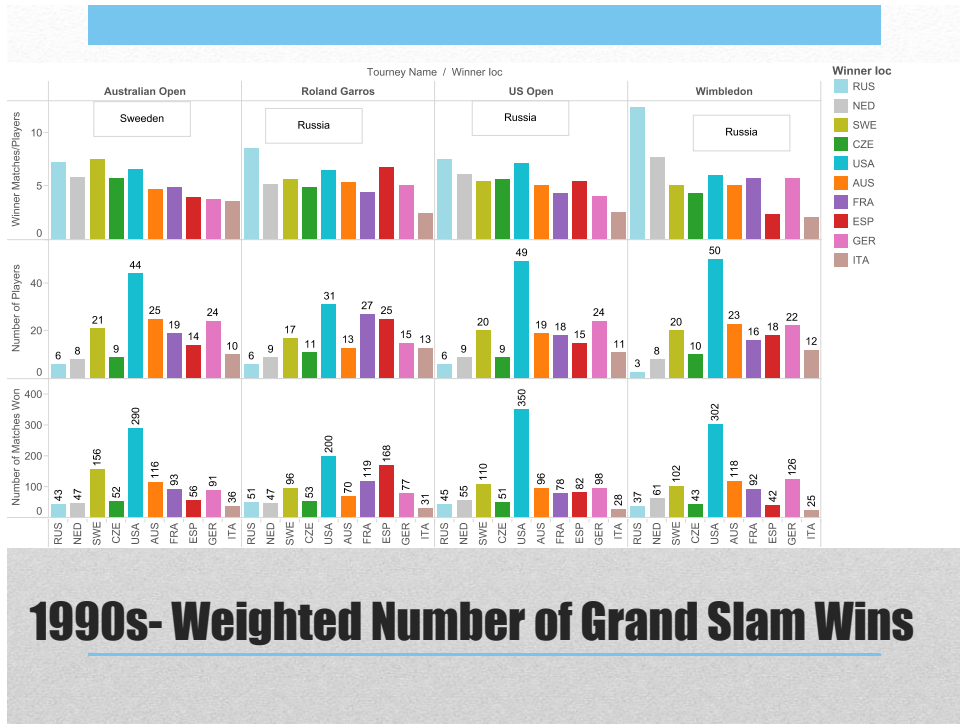




GRAND SLAM WINS

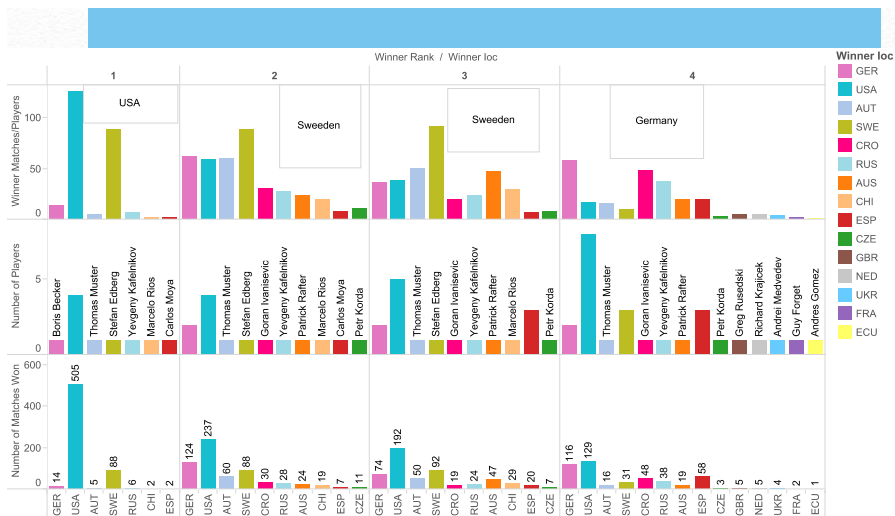
1970-2000



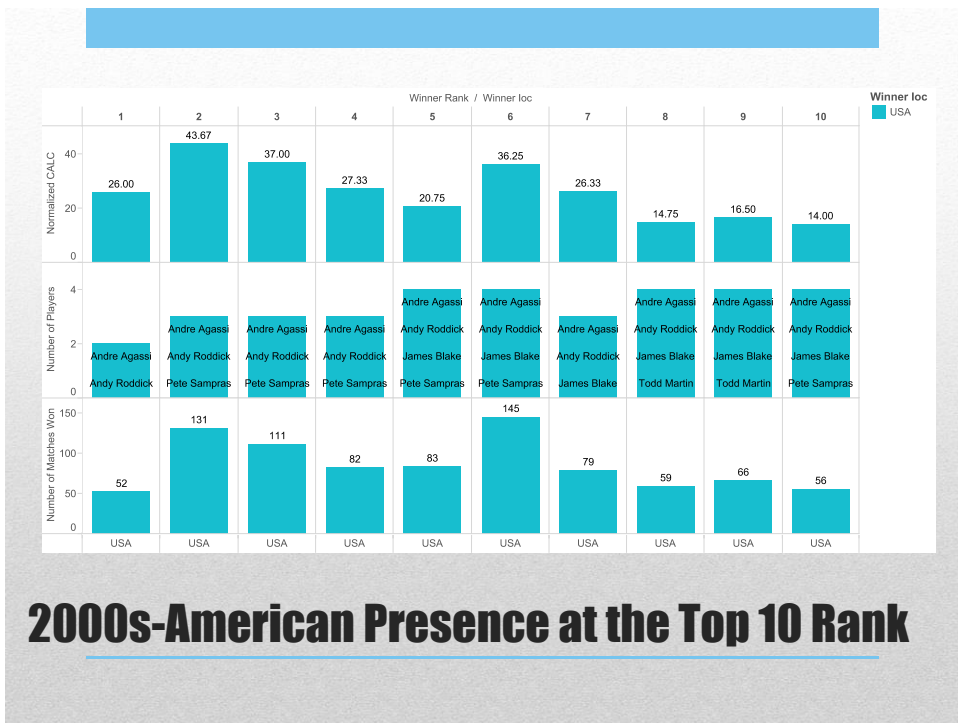
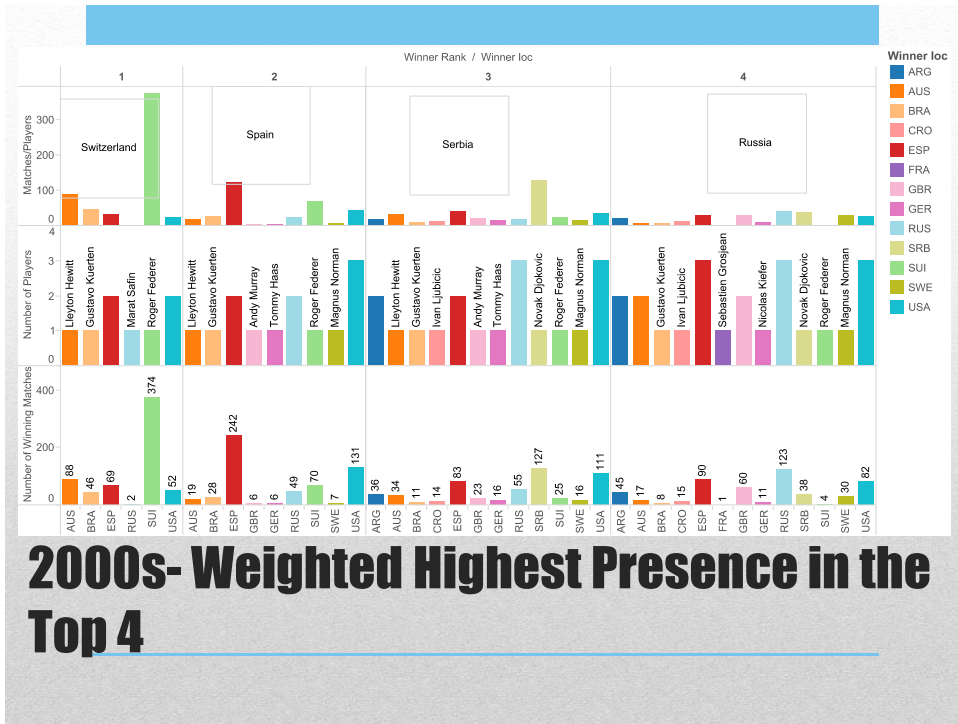


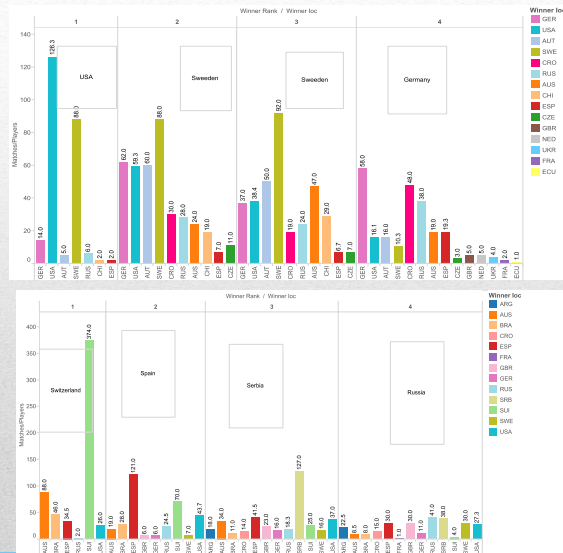
ATP RANK:WHICH COUNTRY COMES OUT ON TOP?

1990-2000 Data only



1990s-Weighted Highest Presence in the Top 4

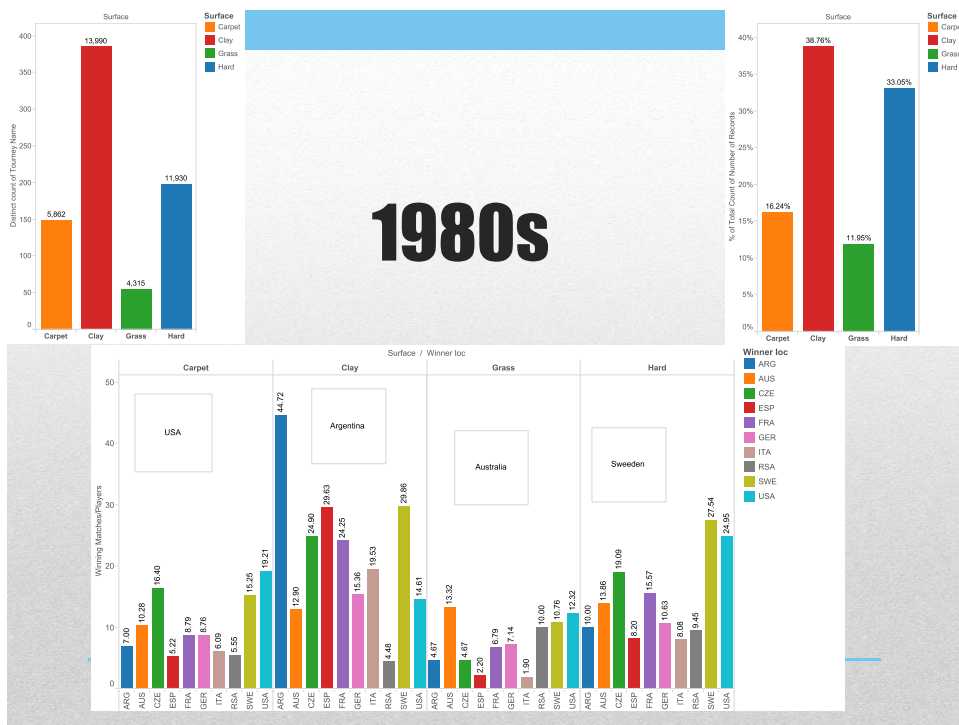
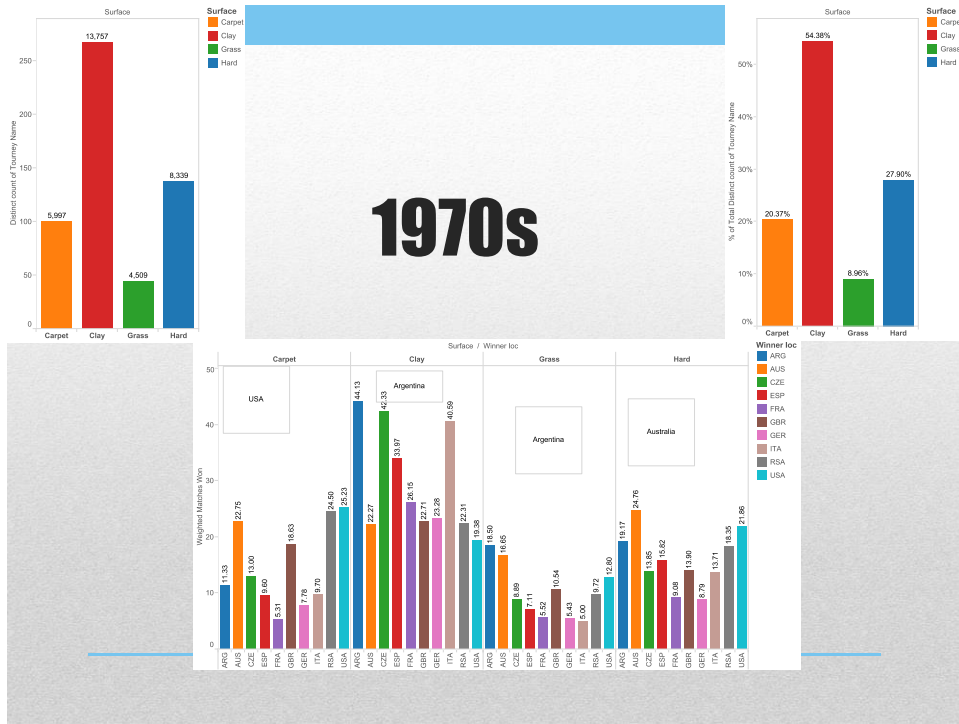


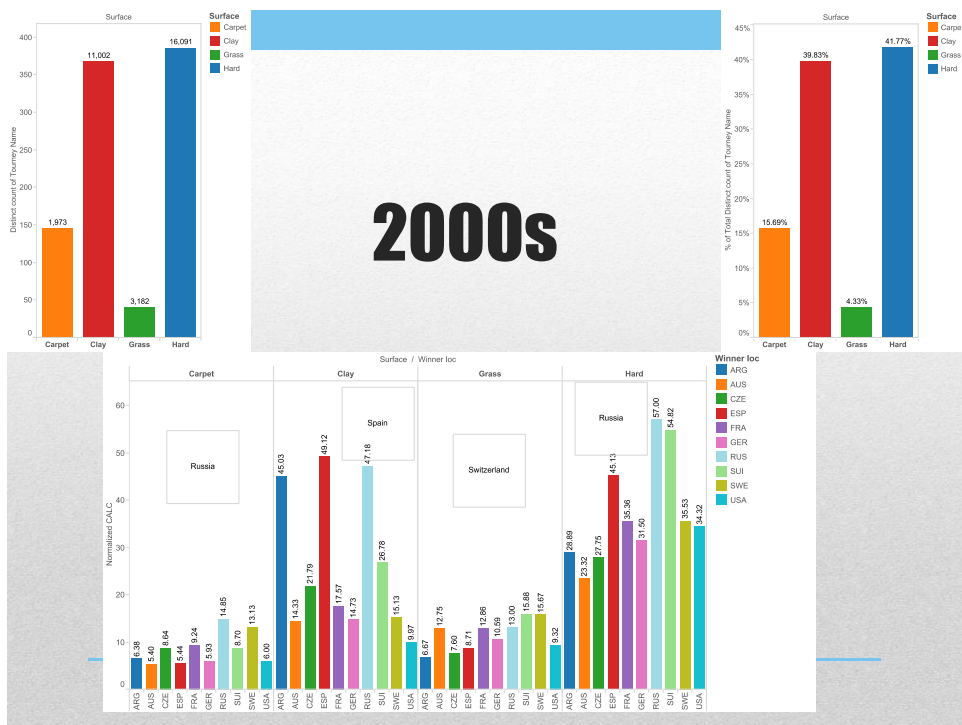
1
9
9
0
s2
0
0
0
s

Two Decade Rank Comparison

WHO DOMINATES ON CLAY? HARDTOP? GRASS?

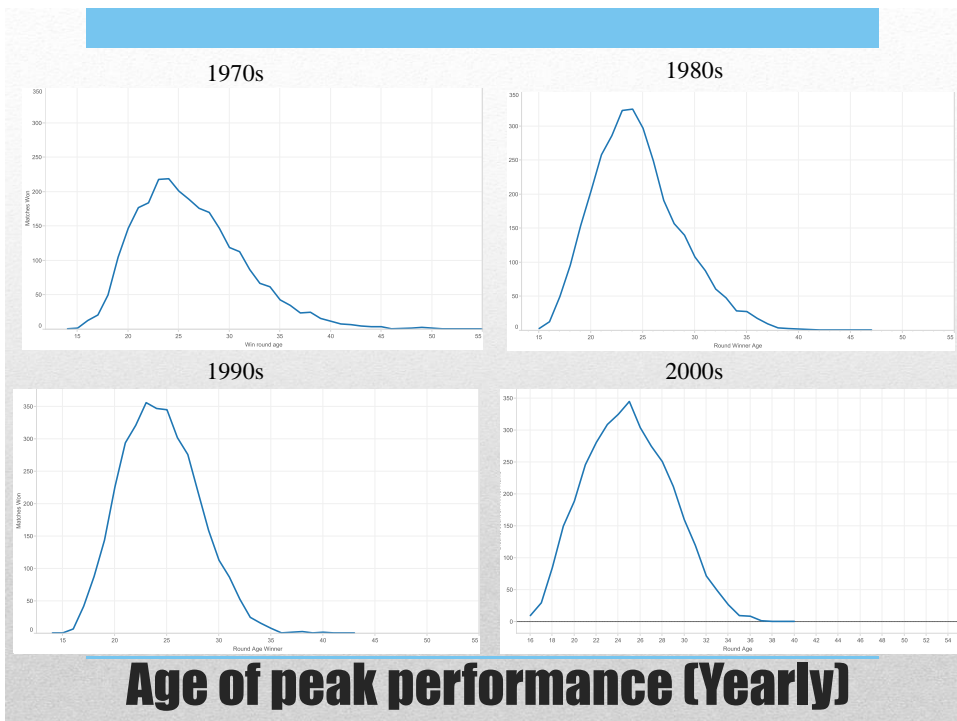
1970-2000





DOES AGE MATTER?

1970-2000



- Earlier decades: More variation in winners of each attribute
- USA:
 - Downward slope for number of players
 - Are not making the top 4 rankings
 - Not winning as many matches
- Future analysis
 - Predictive Analysis- What is in store for 2010s?

Decade	Raw Number of Wins	Number of Players	Weighted Number of Wins	Grand Slams				Top Ranks				Surface			
				Australian Open	French Open	US Open	Wimbledon	1	2	3	4	Carpet	Clay	Grass	Hard
70s	USA	USA	CZE	ARG	ITL	ARG	USA					USA	ARG	ARG	AUS
80s	USA	USA	SWE	SWE	SWE	SWE	SWE					USA	ARG	AUS	SWE
90s	USA	USA	RUS	SWE	RUS	RUS	RUS	USA	SWE	SWE	GER	RUS	ESP	RUS	RUS
00s	ESP	USA	RUS	SUI	SUI	SUI	SUI	SUI	ESP	SER	RUS	RUS	ESP	SUI	RUS

Conclusion and Winning Country Profile