



Data Science

- Intro
- Past – Inspiration
- Present – Getting Practice
- Future



What Is Data Science To Me?

- Applying statistics to data
- Lots of data means data engineering (Petabytes at Chevron)
- Insights → Business Decisions



Past: Why I Chose Data Science

- Many applications across industries/companies
- Uncover new personal interests



Past: Applications

- Finance
 - Most sciences: Environmental, Social, Biological
 - Predict equipment failures
 - Finding oil and gas based on sensors/geology
-

“Where’s the next best place to drill for oil? Or next best customer service to target?”

- Data Analytics: Mark Miller; Chevron Youtube (Sept 2014)



Past: New Interests

“Facilities engineering encompasses all of the engineering disciplines that are associated with the design and construction of facilities required to produce our energy”

- Offering Opportunities for Engineering Graduates, Chevron Youtube

(Feb 2013)



Present: What I'm Doing Now

- Small projects to practice with
- Super Rugby Database

Present: Super Rugby DB

SRDB

Home
Seasons
Records
Descriptive
Totals
Draws

Super Rugby DB

Team	Country	Games	Wins	Win Tally	For	Rank	O. Rank	D. Rank	Titles
Blues	NZL [1996]	311	162	15	8269	5	3	8	3
Brumbies	AUS [1996]	317	181	48	8286	3	6	2	2
Bulls	RSA [1996]	311	154	-6	8058	10	7	12	3
Cheetahs	RSA [1997]	187	61	-67	4344	14	13	15	0
Chiefs	NZL [1996]	314	174	38	8215	6	5	7	2
Crusaders	NZL [1996]	334	234	129	10040	1	1	1	9
Western Force	AUS [2006]	175	60	-57	3399	14	18	10	0
Highlanders	NZL [1996]	313	157	2	7762	9	9	9	1
Hurricanes	NZL [1996]	315	180	42	8504	3	2	6	1
Jaguars	ARG [2016]	47	20	-7	1212	11	8	13	0
Southern Kings	RSA [2013]	48	12	-22	1013	18	17	17	0
Lions	RSA [1996]	297	114	-77	7233	12	11	14	0
Rebels	AUS [2011]	126	38	-48	2688	16	16	16	0
Reds	AUS [1996]	307	140	-27	6899	12	14	11	1
Sharks	RSA [1996]	318	166	21	7839	7	10	5	0
Stormers	RSA [1996]	298	160	20	7031	8	12	4	0
Sunwolves	JPN [2016]	46	7	-34	1012	17	15	18	0
Waratahs	AUS [1996]	312	168	30	8178	2	4	3	1

Present: Super Rugby DB

Crusaders (NZL)

Averages

	Games	Record	High	Low	Avg	AvgDiff	MedDiff	OR	DR	Net
CRU	334	234-100 [70%]	96	0	30.0	9.0	7.0	0.62	0.27	0.89

Game Log

Home	Score	Away	Date	Location	Winner	
Crusaders	37 - 18	Lions	2018-08-04	Rugby League Park	Crusaders	[vs LIO] [loc]
Crusaders	30 - 12	Hurricanes	2018-07-28	Rugby League Park	Crusaders	[vs HUR] [loc]
Crusaders	40 - 10	Sharks	2018-07-21	Rugby League Park	Crusaders	[vs SHA] [loc]
Crusaders	54 - 17	Blues	2018-07-14	Rugby League Park	Crusaders	[vs BLU] [loc]
Crusaders	45 - 22	Highlanders	2018-07-06	Rugby League Park	Crusaders	[vs HIG] [loc]
Chiefs	20 - 34	Crusaders	2018-06-02	Waikato Stadium	Crusaders	[vs CHI] [loc]



Present: Super Rugby DB

Descriptive Statistics

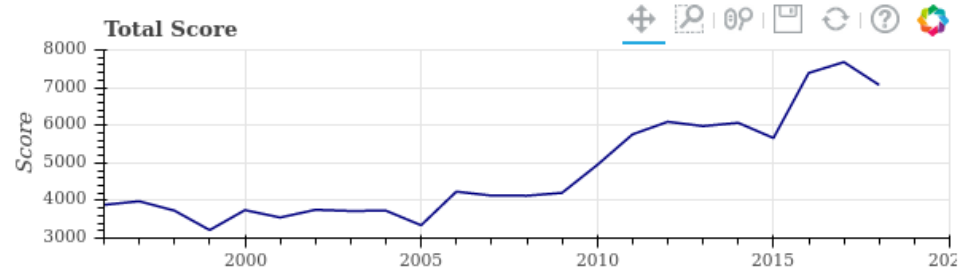
All Time Scores

Team	Mean		Median		Max		Min	
	@home	@away	@home	@away	@home	@away	@home	@away
Blues	28.7	24.41	26.0	23.0	74.0	60.0	5.0	3.0
Brumbies	30.49	21.7	28.0	21.0	73.0	52.0	3.0	0.0
Bulls	30.77	20.95	29.0	20.0	92.0	60.0	7.0	0.0
Cheetahs	26.52	19.9	26.5	20.0	92.0	53.0	6.0	0.0
Chiefs	27.88	25.81	26.0	21.0	61.0	73.0	6.0	3.0

Present: Super Rugby DB

2015	Hurricanes	Highlanders	250	5660	22.64	0.47	0.14	58	1.2
2016	Hurricanes	Lions	284	7389	26.02	0.54	-0.2	92	1.91
2017	Lions	Crusaders	284	7678	27.04	0.56	-0.3	94	1.95
2018	Crusaders	Lions	254	7069	27.83	0.58	-0.38	77	1.59

Graphs





Present: Super Rugby DB

- Simple Data Engineering (Wikipedia → database)
- Working with limited stats knowledge
- Working with limited data
- Balance Information and Brevity



Future: How I Will Improve

- Using what I know
- School → Insights → Business Decisions



End

Thank You :)