

# **The Power of Proximity**

Using Network Science to Analyze the Link Between Coaching  
Communities and Team Success in the National Football League

Atul Venkatesh, Sid Singh, Charlie Betts

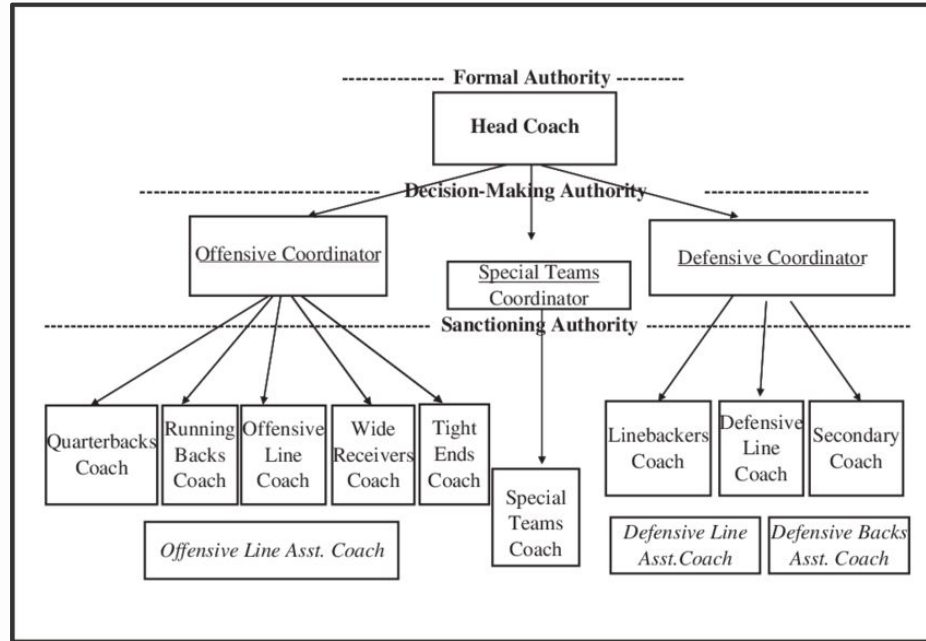
**01**

# **Setting the Scene**

# A Recent Test Case



# Coaching Staff Hierarchy



**“NFL coaches clearly affect points scored** and the point margin.

The estimates imply that coaches explain 18 to 25 percent of within-team, between-season variation in points allowed and point margins...They significantly affect the number of fumbles and penalties a team commits.”

- Christopher R. Berry and Anthony Fowler, *University of Chicago*

**02**

# **Data Collection**

# Data Collection Methods

## 1. Who?

Head Coaches,  
Coordinators, Positional  
Coaches and Quality  
Control Coaches from  
**2010-2024**

## 2. How?

Coaching Data:  
*pro-football-history.com*.  
Standings Data: *NFLFastR*.

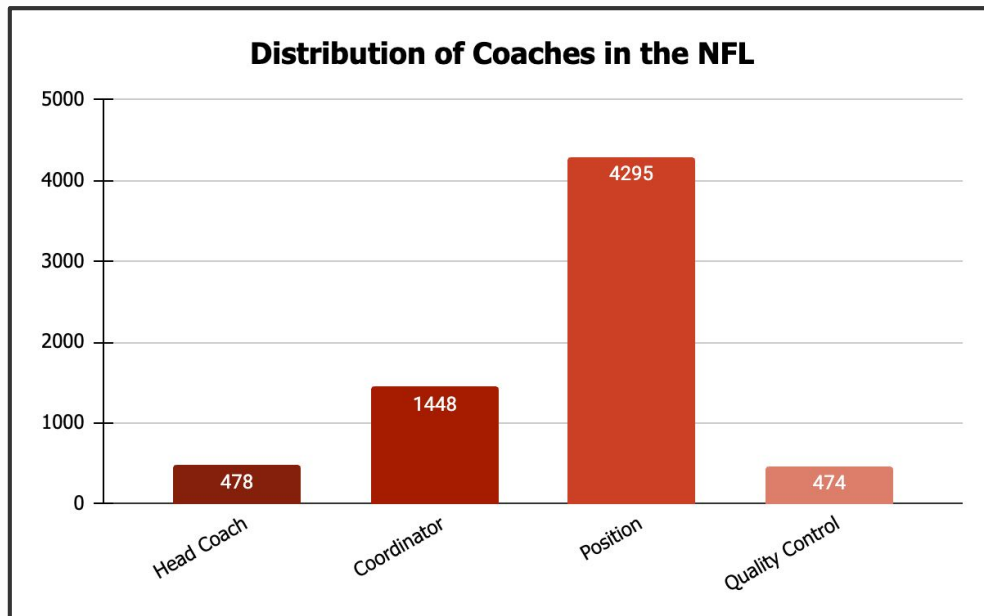
## 3. Why?

Decision-Making  
Influence and Coaching  
Ladder

# Summary Statistics

**6695** total coaches

- **478** head coaches
- **1448** coordinators
- **4295** positional coaches
- **474** quality control coaches





**03**

# **Data Wrangling**

# Weighting Coaching Positions

10 → Head Coaches

7.5 → Coordinators

5 → Positional Coaches

2.5 → Quality Control Coaches

# Determining Connections

6695 coaches → **44435 connections**

Strength of Connection (SoC): Sum of Coaching Weights

For example:

→ Head Coach — Coordinator:  $10 + 7.5 = 17.5$

→ Coordinator — Positional Coach:  $7.5 + 5 = 12.5$

→ Positional Coach — Quality Control Coach:  $5 + 2.5 = 7.5$

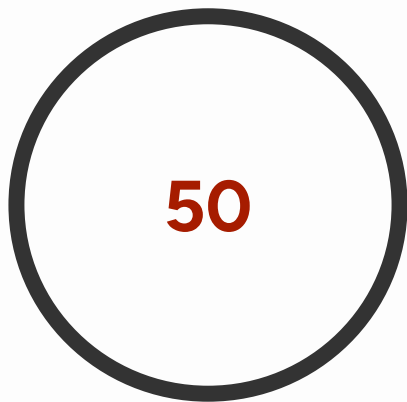
Sum of SoC for every year two coaches coached together =

**Aggregated Closeness Score**

**04**

# **Clustering Analysis**

# Pruning

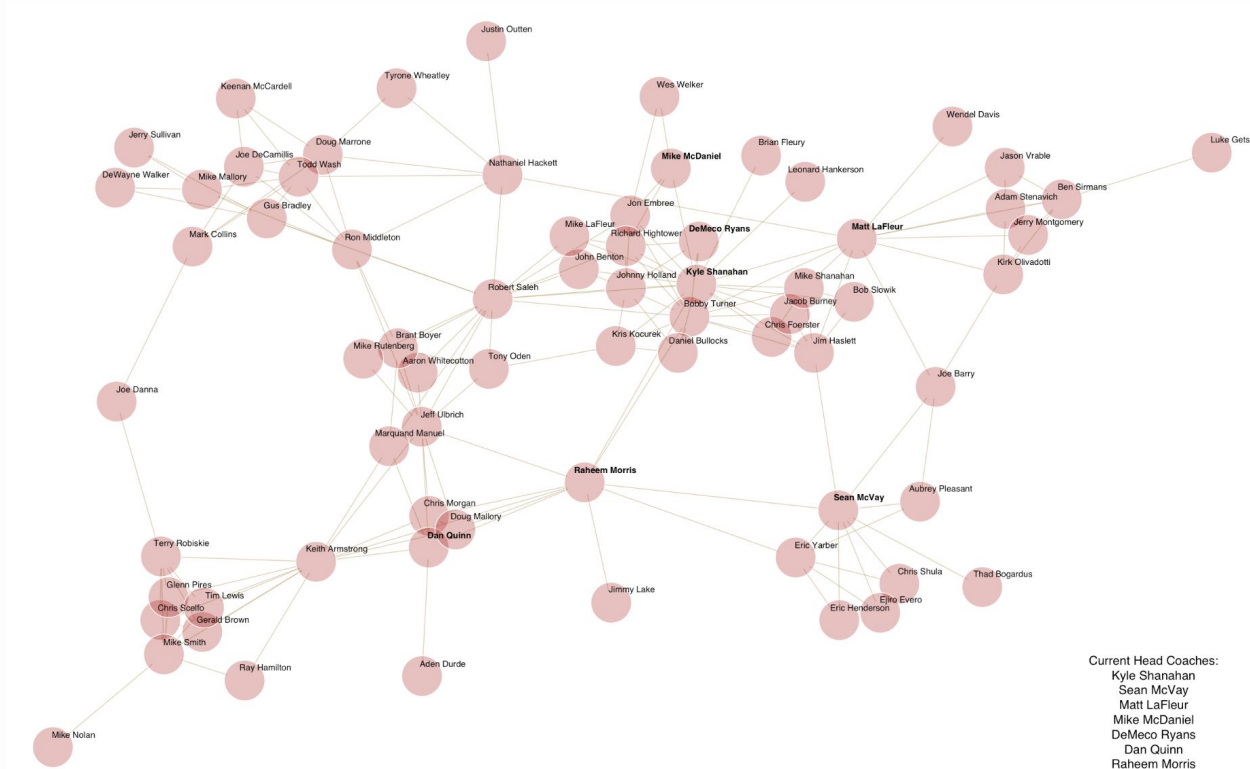


Connection Strength **Threshold**

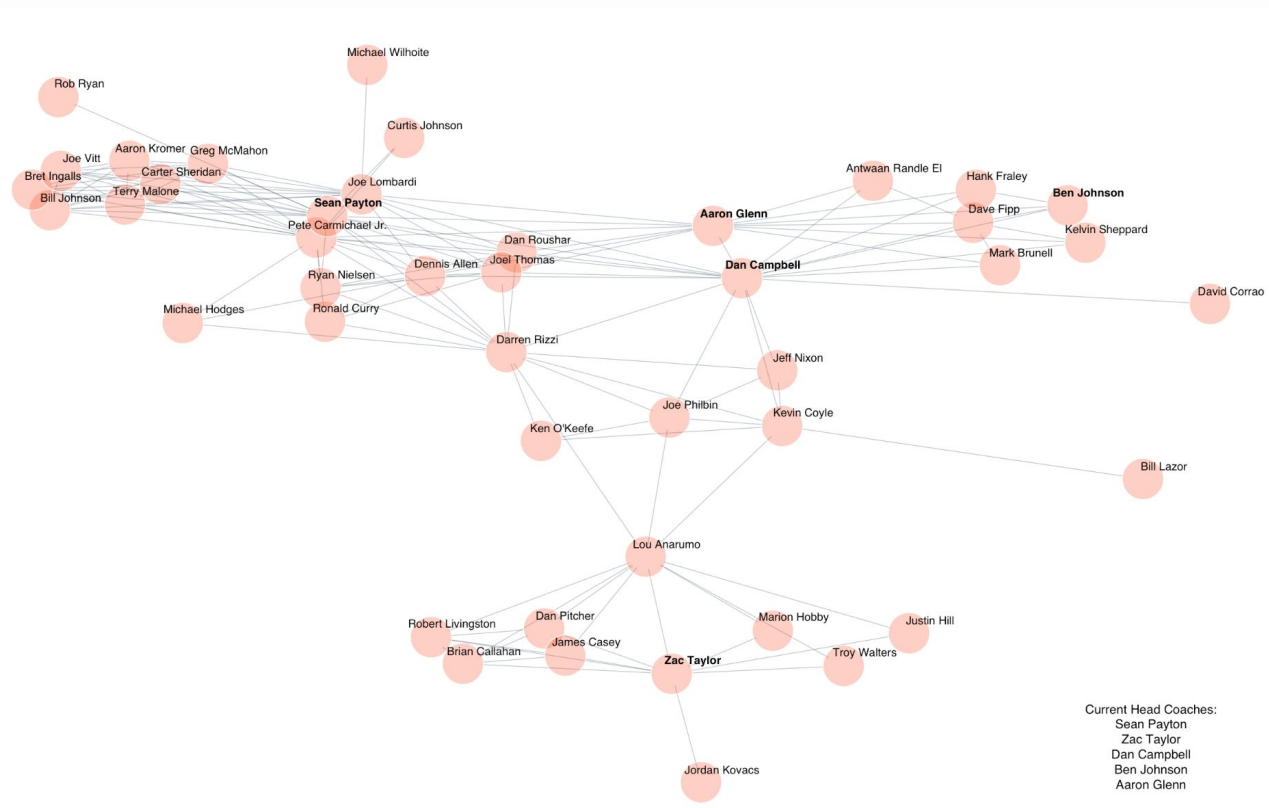
# Conducting the Clustering Algorithm

- Conducted a fast-greedy modularity maximizing clustering algorithm
- The algorithm split the network into 19 distinct communities
- 13 were determined to be significant (have several big names)
- Each significant community was named after a coach we know had influence on their edges

# The Shanahan Community

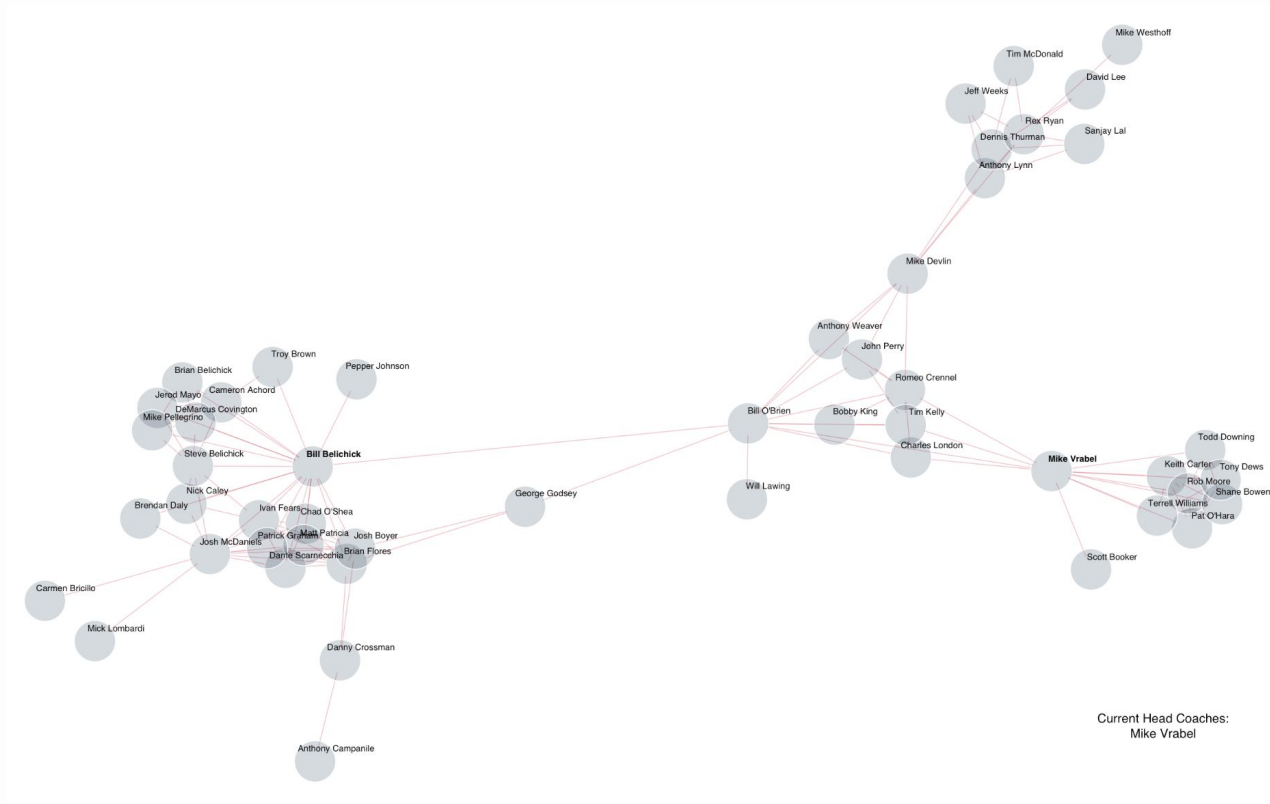


# The Payton Community





# The Belichick Community



# Other Communities

- **Andy Reid Community** - Includes Nick Sirianni + Jonathan Gannon
- **Pete Carroll Community** - Includes Dave Canales + Brian Schottenheimer
- **Ron Rivera Community** - Includes Sean McDermott + Brian Daboll
- **Mike Tomlin Community** - No active head coach
- **John Harbaugh Community** - Includes Mike Macdonald
- **Vic Fangio Community** - Includes Jim Harbaugh
- **Jason Garrett/Jon Gruden Community** - No active head coach
- **Marvin Lewis/Mike Zimmer Community** - Includes Kevin O'Connell + Kevin Stefanski
- **Mike McCarthy Community** - Includes Kellen Moore
- **Bruce Arians Community** - Includes Todd Bowles

**05**

# **“Six Degrees of Kyle Shanahan”**

# Motivation and Methods

How much *influence* does Kyle Shanahan have in the network?

Distance Calculation:

$1/(\text{Aggregated closeness score}) + 1/(\text{years coached})$

Conducted an algorithm to determine the *shortest path* from every coach to Shanahan.

Bacon with a twist



# Results

How close are major Head Coaches to Kyle Shanahan?			
	Coach Name	Path	Score
1	Matt LaFleur	Kyle Shanahan - Matt LaFleur	0.18
2	DeMeco Ryans	Kyle Shanahan - DeMeco Ryans	0.213
3	Mike McDaniel	Kyle Shanahan - Mike McDaniel	0.213
4	Raheem Morris	Kyle Shanahan - Raheem Morris	0.27
5	Sean McVay	Kyle Shanahan - Bobby Turner - Sean McVay	0.35
6	Mike Tomlin	Kyle Shanahan - Danny Smith - Mike Tomlin	0.444
7	Pete Carroll	Kyle Shanahan - Brian Schneider - Pete Carroll	0.448
8	Shane Steichen	Kyle Shanahan - Robert Saleh - Gus Bradley - Shane Steichen	0.658
9	Kevin O'Connell	Kyle Shanahan - Bobby Turner - Sean McVay - Wes Phillips - Kevin O'Connell	0.682
10	Sean McDermott	Kyle Shanahan - Matt LaFleur - Ben Sirmans - Rob Boras - Sean McDermott	0.765
11	Dan Campbell	Kyle Shanahan - John Benton - Darren Rizzi - Dan Campbell	0.781
12	Mike Vrabel	Kyle Shanahan - Jim Haslett - Mike Vrabel	0.8
13	Andy Reid	Kyle Shanahan - Jon Embree - Joe Cullen - Andy Heck - Andy Reid	0.817
14	Mike Macdonald	Kyle Shanahan - Jon Embree - Joe Cullen - Mike Macdonald	0.82

**06**

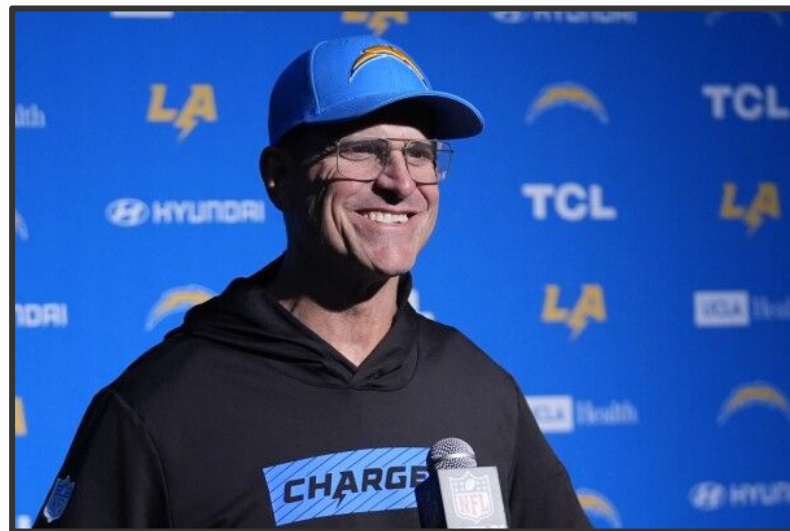
# **Predicting Team Success**

# Motivation

Is the *Harbaugh effect* a global phenomenon?

In other words...

Does the **closeness of coaching staff** have a significant effect on team performance?



# Methods

$$\text{Team Closeness Score} = \frac{1}{N} \sum_{i=1}^N (\text{Aggregated Closeness Score}_i + \text{Years Coached Together}_i)$$

Conducted 3 Regression Models

**Predictor** = Team Closeness Score

**Controls** = Strength of Schedule and Win Percentage Previous Year

	Method	Predictor	Outcome
Regression 1	Linear	Team Closeness Score	Win percentage
Regression 2	Logistic	Team Closeness Score	Made Playoffs (yes/no)
Regression 3	Logistic	Team Closeness Score	Won Division (yes/no)



# Results

Table 1: Effect of coaching staff closeness on team win percentage

	Estimate	Std. Error	t Value	Pr(>  t )
<b>Coefficients</b>				
Intercept	1.6557	0.103874	15.940	<2.0e-16***
Closeness Score	0.0038	0.0004558	8.402	5.15e-13***
Lag Win %	0.2309	0.0369745	6.245	9.43e-10***
Strength of Schedule	-2.8151	0.2014151	-13.977	<2.0e-16***

Notes: \*\*\*/\*\*/\* denotes significance at the 5/1/.1 percent.  
Residual standard error: 0.1392 on 474 degrees of freedom  
Multiple R-squared: 0.473, Adjusted R-squared: 0.4697  
F-statistic: 141.8 on 3 and 474 DF, p-value: < 2e-16

Table 2: Effect of coaching staff closeness on team playoff probability

	Estimate	Std. Error	z Value	Pr(>  z )
<b>Coefficients</b>				
Intercept	13.3461	2.0616	6.474	< 2.0 × 10 <sup>-16</sup> ***
Closeness Score	0.0525	0.0092	5.726	1.03 × 10 <sup>-7</sup> ***
Lag Win %	2.3590	0.6689	3.527	4.21e-4***
Strength of Schedule	-33.9313	4.3035	-7.885	< 2.0 × 10 <sup>-16</sup> ***

Notes: \*\*\*/\*\*/\* denotes significance at the 5/1/.1 percent.  
Null deviance: 642.41 on 477 degrees of freedom  
Residual deviance: 472.73 on 474 degrees of freedom  
AIC: 480.73

Interpreting table 2:  
An increase of 1 in  
*Team Closeness*  
Score leads to a  
**1.05% Increase** in  
*Playoff Probability*

Table 3: Effect of coaching staff closeness on division winning probability

	Estimate	Std. Error	z Value	Pr(>  z )
<b>Coefficients</b>				
Intercept	12.9996	2.2953	5.664	1.48 × 10 <sup>-8</sup> ***
Closeness Score	0.0418	0.0088	4.763	1.91 × 10 <sup>-6</sup> ***
Lag Win %	3.4843	0.7792	4.472	7.76 × 10 <sup>-7</sup> ***
Strength of Schedule	-35.6366	4.8895	-7.288	3.14 × 10 <sup>-11</sup> ***

Notes: \*\*\*/\*\*/\* denotes significance at the 5/1/.1 percent.  
Null deviance: 538.69 on 477 degrees of freedom  
Residual deviance: 389.81 on 474 degrees of freedom  
AIC: 397.81

Interpreting table 3:  
An increase of 1 in  
*Team Closeness*  
Score leads to a  
**1.04% Increase** in  
*Division Winning*  
*Probability*

**07**

# **Conclusions and Next Steps**

# Conclusions

1

Franchises might benefit from letting connections develop between staff.

2

An applicant's prior connections to existing coaches should be considered during hiring decisions.

3

Teams might benefit from letting head coaches have input into coaching staff additions

4

Coaching chemistry plays a significant role in team success.

# Limitations

- 1** Pre-2010 data could impact connection score/clustering
- 2** Coach value across a position might not be the same
- 3** Confounding variables could impact team success

# Future Research

- 1** Investigate hiring preferences
- 2** Conduct analyses regarding promotions and firings
- 3** Incorporate college coaching networks

# Thanks For Listening!

