NFL Gravity Metric

Authors: Anika Soitkar, Ayush Varma, Mateo Shelton, Michael Zheng, Moulik Chatterjee, Vardaan Sinha

Research Question

How can we evaluate which pass rushers absorb the most attention from the opposition?

Gravity = Actual Attention - Expected Attention

Data: Big Data Bowl 2023 via Kaggle, using game, play, and scouting datasets.

Defining Pass Rush Attention Scores

We calculate attention scores to quantify how much attention a pass rusher receives on each play.

Remember that observed attention will be used to calculate gravity:

Gravity = Observed Attention - Expected Attention

Summary of our process for calculating attention:

1. Filter to only include pass plays that are not screens, RPOs, spikes, or designed rollouts (true pass sets).
2. Filter to only include frames that are in our defined pass rush window (will explain later)
3. For each frame in the pass rush window, calculate a frame-level attention score for each rusher.
4. Aggregate by play to get each rusher’s attention score on each snap.

How We Calculated *Observed* Attention:

Each rusher's **frame-level** attention score is the number of blockers they are occupying at frame x.

At each frame, we define each blocker's "matchup" using a **weighted distance** metric:

- **Weighted Distance = Actual Distance / cos(ϴ)**

* ϴ is the angle between:
- the vector from the blocker to the rusher
- the blocker's orientation direction

Whichever rusher has the smallest positive **weighted distance** value to that blocker is "occupying" that blocker at frame x.

Example:

Each rusher's attention score A(x) at frame x:

- #96: A(x) = 1
- #92: A(x) = 1
- #98: A(x) = 1
- #93: A(x) = 2

Then, aggregate at the **play-level** to get each rusher's attention score **for each play**:

$$\text{Attention} = \frac{1}{N}\sum_{x} A(x)$$

***N is the number of frames in the pass rush window for that play**

Calculating *Expected* Attention: Model Design

GNN	XGBoost
All players influence one another spatially → GNN allows for better interaction-based modeling.	Simple, effective model for multiple engineered features → able to provide a variety of inputs from our data engineering

Potential Ensemble Model → Combining GNN + XGBoost predictions to improve prediction accuracy.

GNN Model Architecture

01 Per-Frame Graph Construction

PyG Data object

Nodes
COMPONENT

N = 22 · 40-D feature

- **centered** x, y, s, a
- **sin / cos** of o, dir
- **team** / role / QB flags
- **height**, weight
- **one-hot** position
- **one-hot** pff_role

Edges
COMPONENT

fully connected · E = 462 · 9-D

- **dx, dy**, distance
- **exp**(-d / 3)
- **same-team** flag
- **off ↔ def** pair flag
- **blocker-rusher** candidate
- **sin / cos** of orient. diff

Globals + Target
COMPONENT

42-D play context

- **down**, distance, quarter
- **formation** / coverage 1-hots
- **OL / DL / DB** counts
- **play-action**, score diff
- Target y broadcast to rushers

02 GATv2 Forward Pass

5-stage pipeline

Node Input MLP
STAGE 1

- Linear (40 → 64)
- ReLU
- Linear (64 → 64)

3 × GATv2Conv
STAGE 2

- **4 heads** × 16 dim → 64
- **edge_dim** = 9
- ELU + dropout(0.2)

$$\alpha_{ij} = \mathrm{softmax}\left(a^{\top} \cdot \mathrm{LeakyReLU}\left(W \cdot [h_i \,\|\, h_j \,\|\, e_{ij}]\right)\right)$$

Global MLP
STAGE 3

- Linear (42 → 64)
- broadcast u[batch]
- to every node

Concat + Head
STAGE 4

- concat (node ‖ global)
- → 128-D
- Linear (128 → 64) + ReLU
- Dropout → Linear (64 → 1)

Per-Node Scalar
STAGE 5

- $\hat{y}_i$ for every node
- supervised on rusher-mask only
- Loss = **masked MSE**
- vs. play-level y

03 Aggregation, Gravity Residual & Outputs

play-rusher rollup → csv

Aggregate to play-rusher
STEP

- **mean** ŷ over the 6 frames
- for each (gameId, playId, rusher_nflId)
- → expected_attention_gnn

Compute gravity
STEP

```
gravity_gnn =
  actual_attention_gnn
  - expected_attention_gnn
```

- z-score & percentile-rank
- over ≥ 25-play sample

gnn_attention_scores
OUTPUT · CSV

- one row per (play, rusher)
- **actual** / expected / gravity

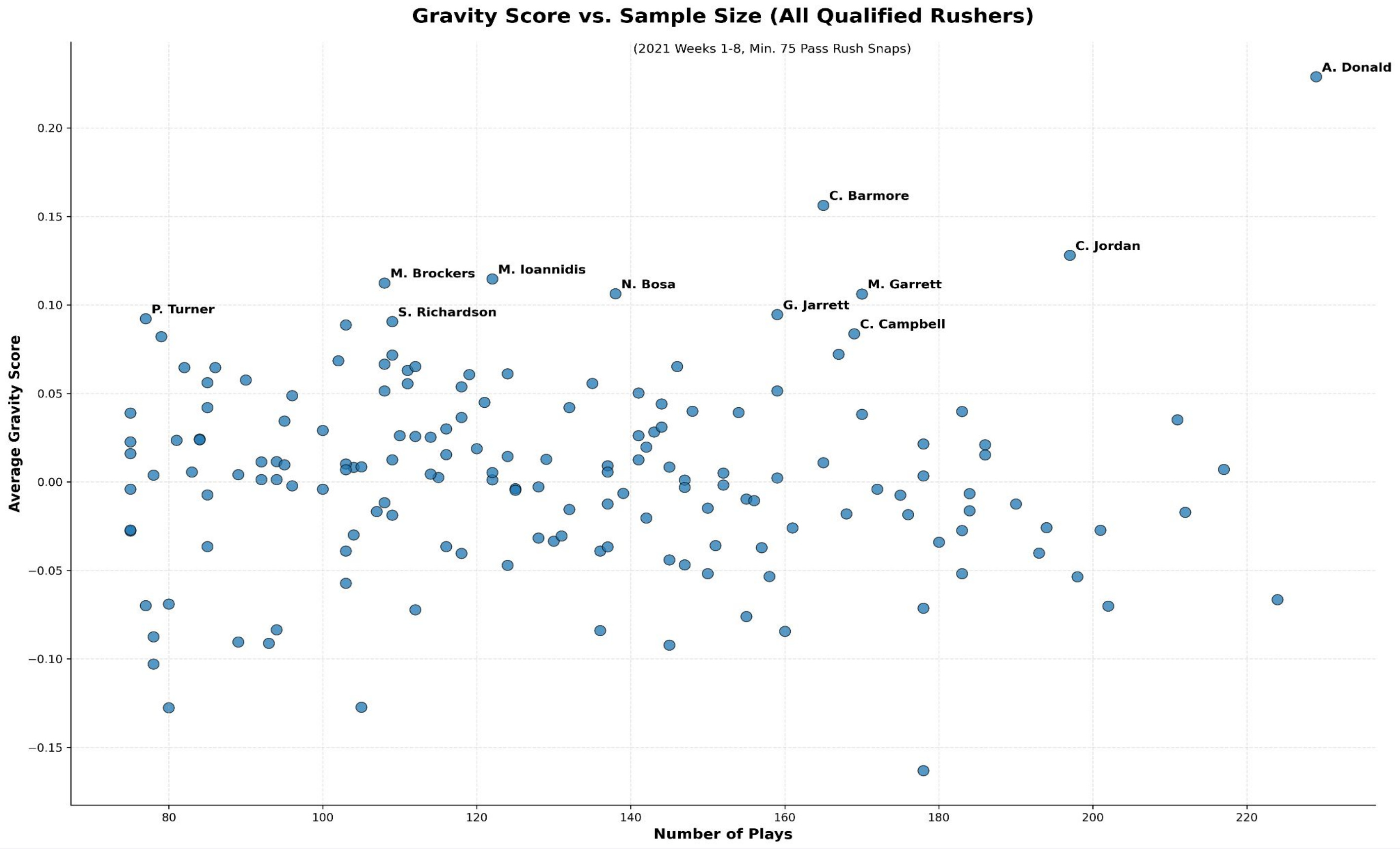
gnn_player_gravity
OUTPUT · CSV

- **qualified-player** rollup
- mean_gravity, gravity_z,
- gravity_pct, position_group

Model Results

XGBoost	
R^2	0.4015
RMSE	0.3813
Most Important Features	pos_DI, isExpectedRusher, euclidean_distance

Gravity Findings



Best Interior Rushers by Gravity and Production

Top 10 Interior Rushers by Gravity and Pass Rush Productivity (PRP)

2021 Weeks 1-8, Min. 75 Pass Rush Snaps | Sorted by Average Z-Score for Gravity and PRP

						Standardized Values		
Name	**Team**	**Position**	**Number of Plays**	**Gravity Score**	**PRP**	**Gravity Z-Score**	**PRP Z-Score**	**Average Z-Score**
Aaron Donald		DI	229	0.229	8.515	3.730	1.593	**2.662**
Christian Barmore		DI	165	0.156	5.758	2.473	0.164	**1.318**
Solomon Thomas		DI	112	0.026	9.375	0.213	2.039	**1.126**
Jonathan Allen		DI	186	0.015	9.677	0.033	2.195	**1.114**
J.J. Watt		DI	167	0.072	7.186	1.017	0.904	**0.961**
Javon Hargrave		DI	150	-0.015	10.000	-0.490	2.363	**0.936**
Star Lotulelei		DI	82	0.065	7.317	0.887	0.972	**0.929**
Chris Jones		DI	122	0.001	9.426	-0.211	2.065	**0.927**
Matt Ioannidis		DI	122	0.115	5.328	1.754	-0.059	**0.847**
Osa Odighizuwa		DI	154	0.039	7.792	0.446	1.218	**0.832**

PRP = (Sacks + ((Hits + Hurries) / 2)) × 100 / Pass Rush Plays (per PFF)

Gravity and PRP z-scores are standardized within position group (DI or ED). Logos from nflverse/ESPN.

Best Edge Rushers by Gravity and Production

Top 10 Edge Rushers by Gravity and Pass Rush Productivity (PRP)

2021 Weeks 1-8, Min. 75 Pass Rush Snaps | Sorted by Average Z-Score for Gravity and PRP

						Standardized Values		
Name	**Team**	**Position**	**Number of Plays**	**Gravity Score**	**PRP**	**Gravity Z-Score**	**PRP Z-Score**	**Average Z-Score**
Myles Garrett		ED	170	0.106	15.588	2.097	2.983	**2.540**
Nick Bosa		ED	138	0.106	13.043	2.101	2.046	**2.074**
Cameron Jordan		ED	197	0.128	6.853	2.523	-0.232	**1.145**
Randy Gregory		ED	129	0.013	12.791	0.285	1.953	**1.119**
Trey Hendrickson		ED	178	0.003	13.202	0.103	2.105	**1.104**
Joey Bosa		ED	159	0.051	9.748	1.035	0.833	**0.934**
Takkarist McKinley		ED	103	0.089	7.282	1.760	-0.075	**0.843**
Micah Parsons		ED	96	-0.002	11.979	-0.007	1.655	**0.824**
T.J. Watt		ED	141	0.026	10.284	0.543	1.030	**0.787**
Jadeveon Clowney		ED	137	0.009	10.949	0.214	1.275	**0.745**

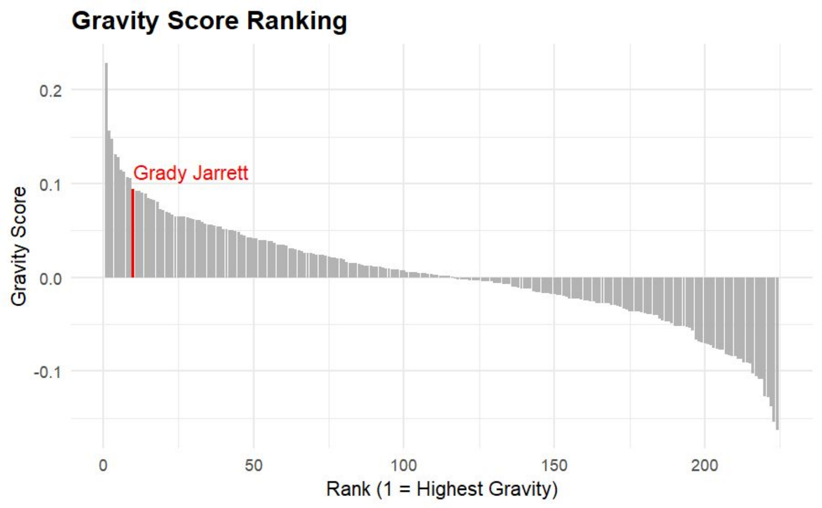
PRP = (Sacks + ((Hits + Hurries) / 2)) × 100 / Pass Rush Plays (per PFF)

Gravity and PRP z-scores are standardized within position group (DI or ED). Logos from nflverse/ESPN.

Case Study: Grady Jarrett

According to PFF, Chicago Bears' Grady Jarrett is widely viewed as a mid-tier interior defensive lineman based on typical metrics.

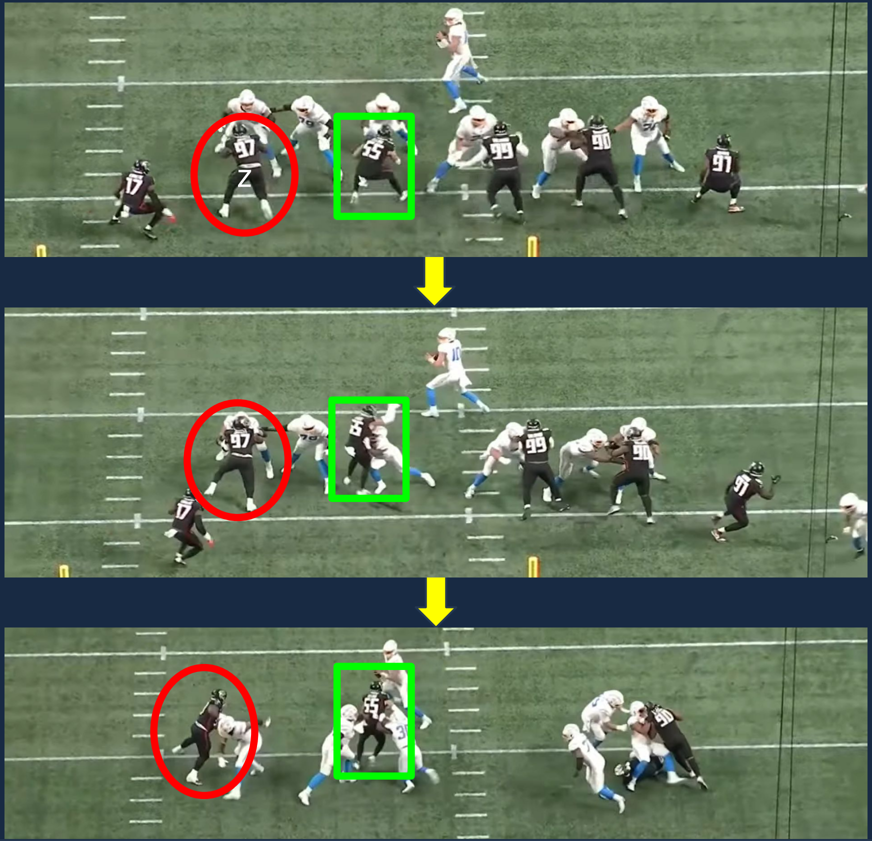
- PFF Overall Grade: 56.7 (76th of 134 IDLs)
- Pass Rush Grade: 64.6 (55th of 134)
- Run Defense Grade: 44.5 (105th of 134)



However, our gravity metric shows that he draws more blockers than expected, indicating that he is treated as more of a threat than statistics show.

Across 159 plays, Jarrett had an average attention of 1.5411, compared to an expected attention of 1.4465.

Case Study: Grady Jarrett

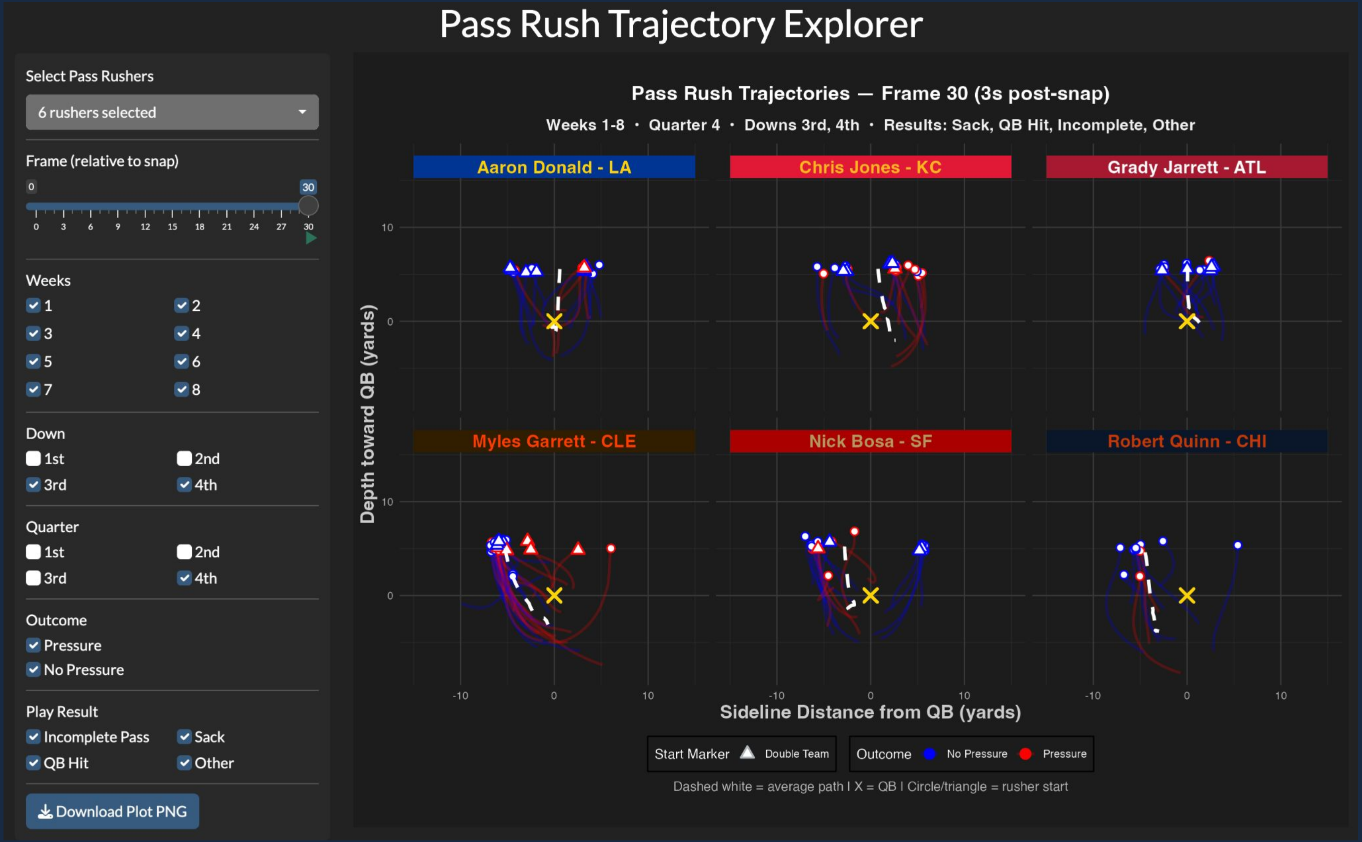


Jarrett is double-teamed immediately after the snap → only 1 blocker on #55, giving him a more favorable path to the QB

Jarrett's double team thins out the protection across the rest of the line.

Even though Jarrett doesn't get the sack, his presence that forces the double team on him creates the opportunity for #55 to get one.

Visualizations



Future Plans + Challenges

1. Finalize and publish our research paper
2. Address assumptions and slight inaccuracies