Winning Strategies: Powering a World Series Victory with Airflow Orchestration

Oliver Dykstra

T
Rangers
MLB
WORLD
SE

Winning Strategies: Powering a World Series Victory with Airflow Orchestration

Oliver Dykstra

Who Am I

Oliver Dykstra

Data Engineer, R&D
Texas Rangers Baseball Club
Joined the club in 2022
odykstra@texasrangers.com

What We’re Talking About

- History of Baseball Statistics
- Big Data Baseball
- The Winning Formula:
  - Upgrading with Airflow
  - Astronomer Take the Wheel
- Putting it all together:

  World Series Victory

A Brief History of Baseball Statistics

That's all baseball is, is numbers; it's run by numbers, averages, percentages and odds...

- Rollie Fingers (Hall of Fame Pitcher)

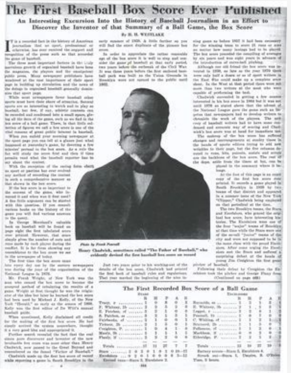
Think Outside the Box(score)

A Brief History of Baseball Statistics

1859 - Henry Chadwick publishes the first Box Score, a set of statistics compiling the runs, hits, outs, assists and errors.

~~ 100 years later ~~

1952 - Topps adds full statistics lines on the back of their annual baseball cards.



The First Baseball Box Score Ever Published

An Interesting Excursion Into the History of Baseball Journalism in an Effort to Discover the Inventor of that Summary of a Ball Game, the Box Score

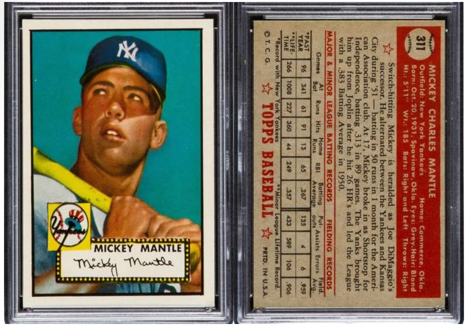
By H. H. WESTLAKE

[illegible]

Henry Chadwick, sometimes called "The Father of Baseball," who evidently devised the first baseball box score on record

The First Recorded Box Score of a Ball Game

[illegible]



311 MICKEY CHARLES MANTLE
Outfield: New York Yankees Home: Commerce, Okla.
Born: Oct. 20, 1931, Spavinaw, Okla. Eyes: Grey. Hair: Blond
Ht.: 5'11" Wt.: 185 Bats: Right and Left Throws: Right

Switch-hitting Mickey is heralded as Joe DiMaggio's successor. He alternated between the Yankees and Kansas City during '51 — batting in 50 runs in 1 month for the American Association club. At 17, Mickey broke in as Shortstop for Independence, batting .313 in 89 games. The Yanks brought him up from Joplin after he hit 26 HR's and led the League with a .383 Batting Average in 1950.

	MAJOR & MINOR LEAGUE BATTING RECORDS							FIELDING RECORDS			
	Games	At Bat	Runs	Hits	Home Runs	RBI	Batting Average	Put-outs	Assists	Errors	Field. Avg.
*PAST YEAR	96	341	61	91	13	65	.267	135	4	6	.959
**LIFE-TIME	266	1008	227	360	44	249	.357	433	589	106	.906

*Record with New York Yankees **Minor League Lifetime Record.

© T. C. G. TOPPS BASEBALL PRTD. IN U.S.A.

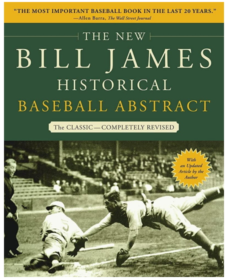
Think Outside the Box(score)

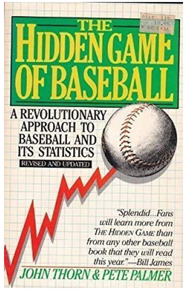
A Brief History of Baseball Statistics

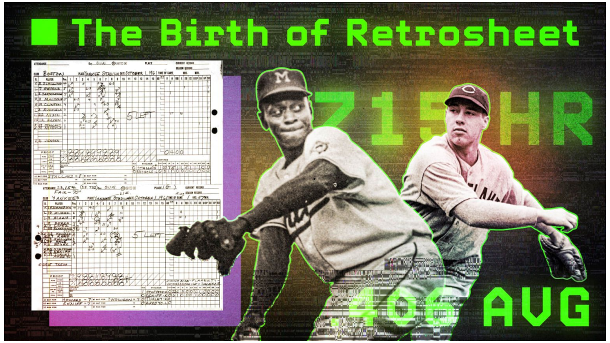
1977 - Bill James publishes his first "Baseball Abstract" which becomes a national bestseller in the early 80s.

1989 - Retrosheet begins massive compilation and online publishing of old box scores and play-by-plays, allowing droves of historical research never before possible.

1996 - Baseball Prospectus begins publication of their annual and website - introduces statistics community to VORP, PECOTA, Pitcher Abuse points and more.


"THE MOST IMPORTANT BASEBALL BOOK IN THE LAST 20 YEARS." —Allen Barra, The Wall Street Journal
THE NEW
BILL JAMES
HISTORICAL
BASEBALL ABSTRACT
The CLASSIC—COMPLETELY REVISED
With an Updated Article by the Author


THE HIDDEN GAME OF BASEBALL
A REVOLUTIONARY APPROACH TO BASEBALL AND ITS STATISTICS
REVISED AND UPDATED
"Splendid...Fans will learn more from The Hidden Game than from any other baseball book that they will read this year."—Bill James
JOHN THORN & PETE PALMER



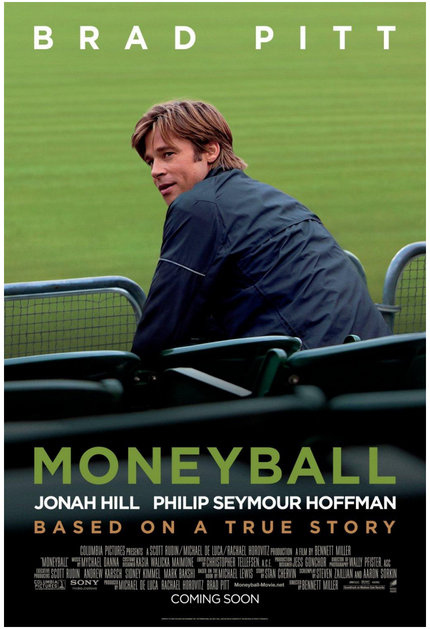
Think Outside the Box(score)

A Brief History of Baseball Statistics

2001 - 2003 - Billy Beane’s Moneyball A’s use data-driven insights to identify market inefficiencies within the game of baseball.

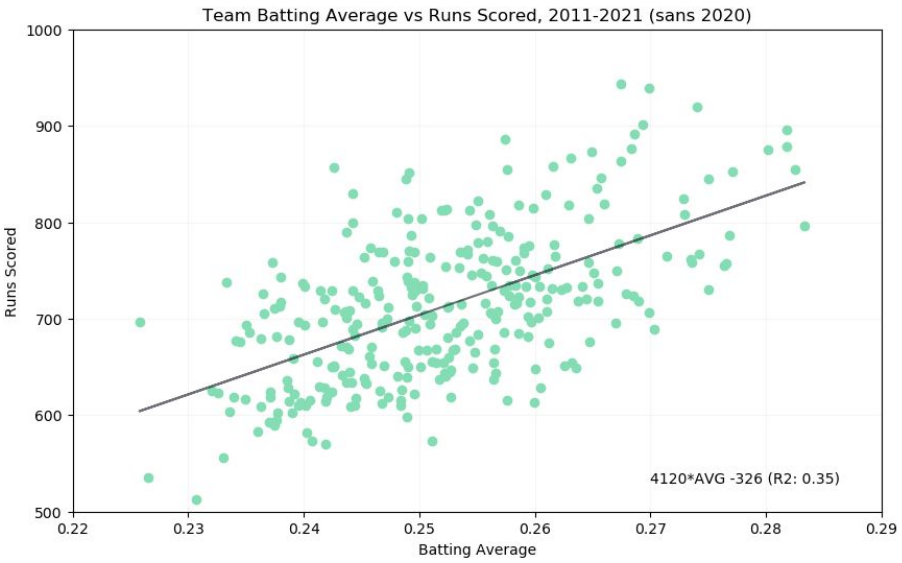
2003 - Michael Lewis publishes his book, “Moneyball: The Art of Winning an Unfair Game” about those A’s.

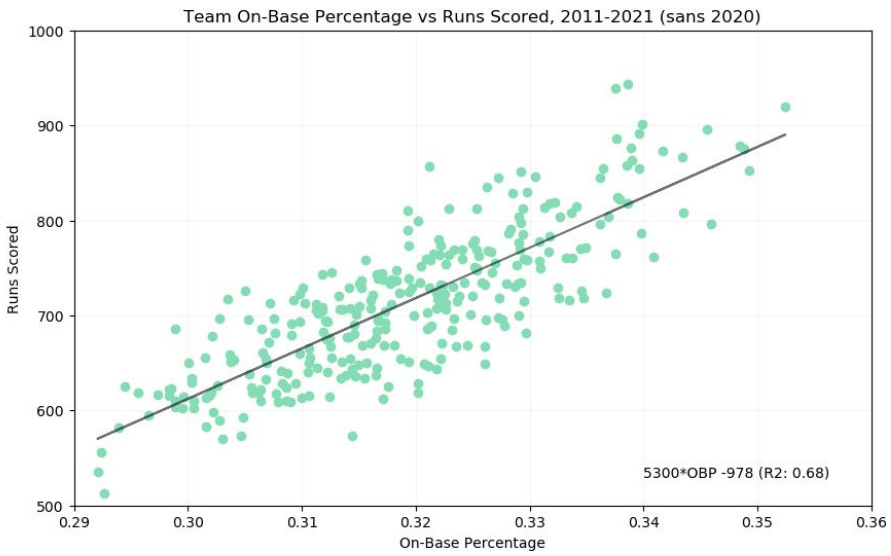
2011 - Brad Pitt stars in the movie adaptation of the same name.



Think Outside the Box(score)

A Brief History of Baseball Statistics





Think Outside the Box(score)

A Brief History of Baseball Statistics

Billy Beane identified a **market inefficiency**.

The market priced players with high batting averages higher than those with high on-base percentages. However, on-base percentage has a **higher correlation** to total runs scored.

The Oakland A's used this information to acquire players undervalued by the market that could help them compete with higher payroll teams.

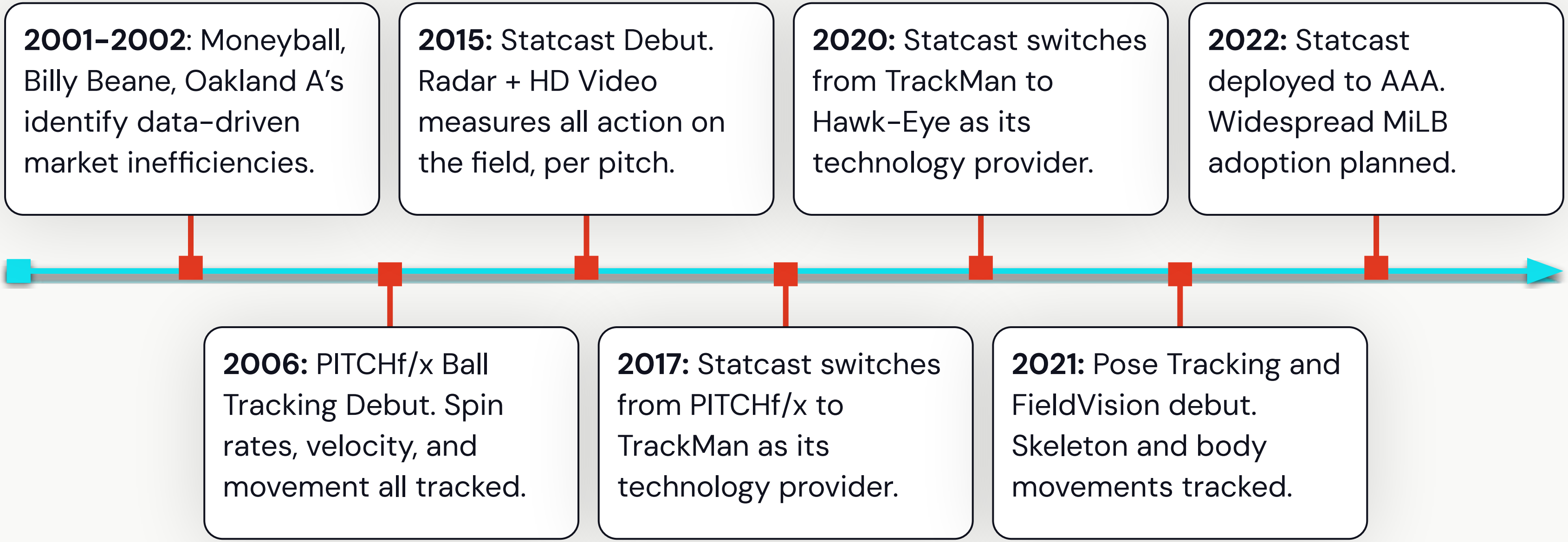
This data-driven decision **disrupted the industry** and left a legacy far beyond baseball.

Big Data Baseball

> “If you challenge conventional wisdom, you will find ways to do things much better than they are currently done.”
>
> – Bill James
>
> (Founder of Sabermetrics)

The Statcast Revolution

Big Data Baseball



2001-2002: Moneyball, Billy Beane, Oakland A’s identify data-driven market inefficiencies.

2006: PITCHf/x Ball Tracking Debut. Spin rates, velocity, and movement all tracked.

2015: Statcast Debut. Radar + HD Video measures all action on the field, per pitch.

2017: Statcast switches from PITCHf/x to TrackMan as its technology provider.

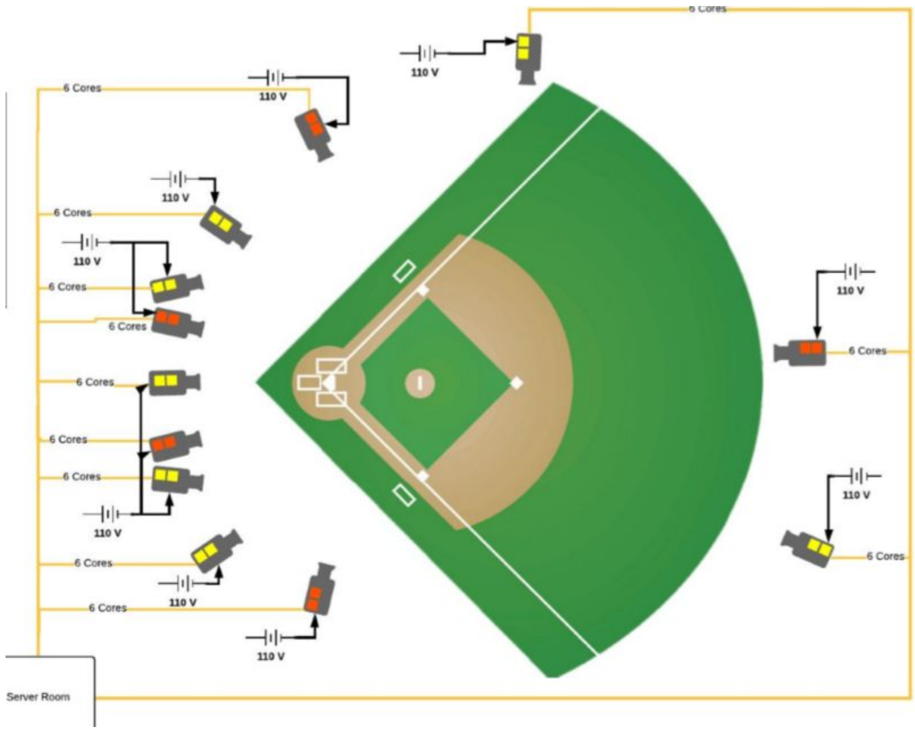
2020: Statcast switches from TrackMan to Hawk-Eye as its technology provider.

2021: Pose Tracking and FieldVision debut. Skeleton and body movements tracked.

2022: Statcast deployed to AAA. Widespread MiLB adoption planned.

The Statcast Revolution

Big Data Baseball

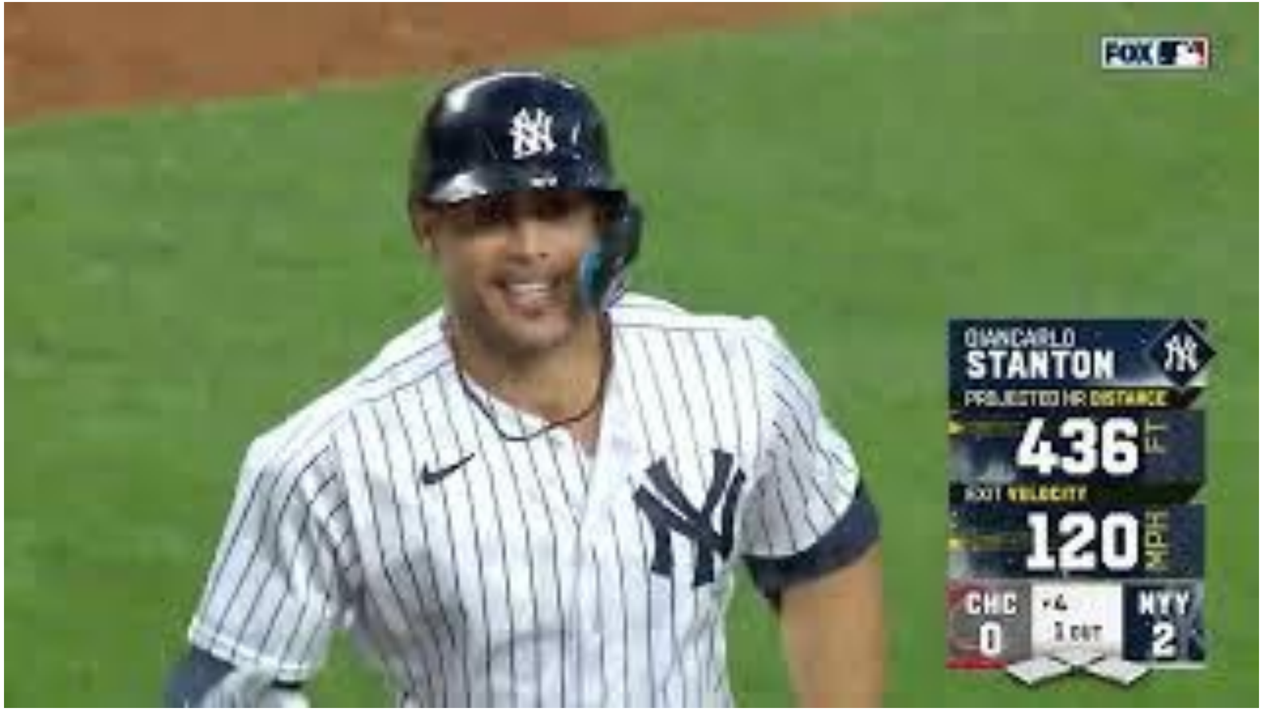


The Statcast Revolution

Big Data Baseball

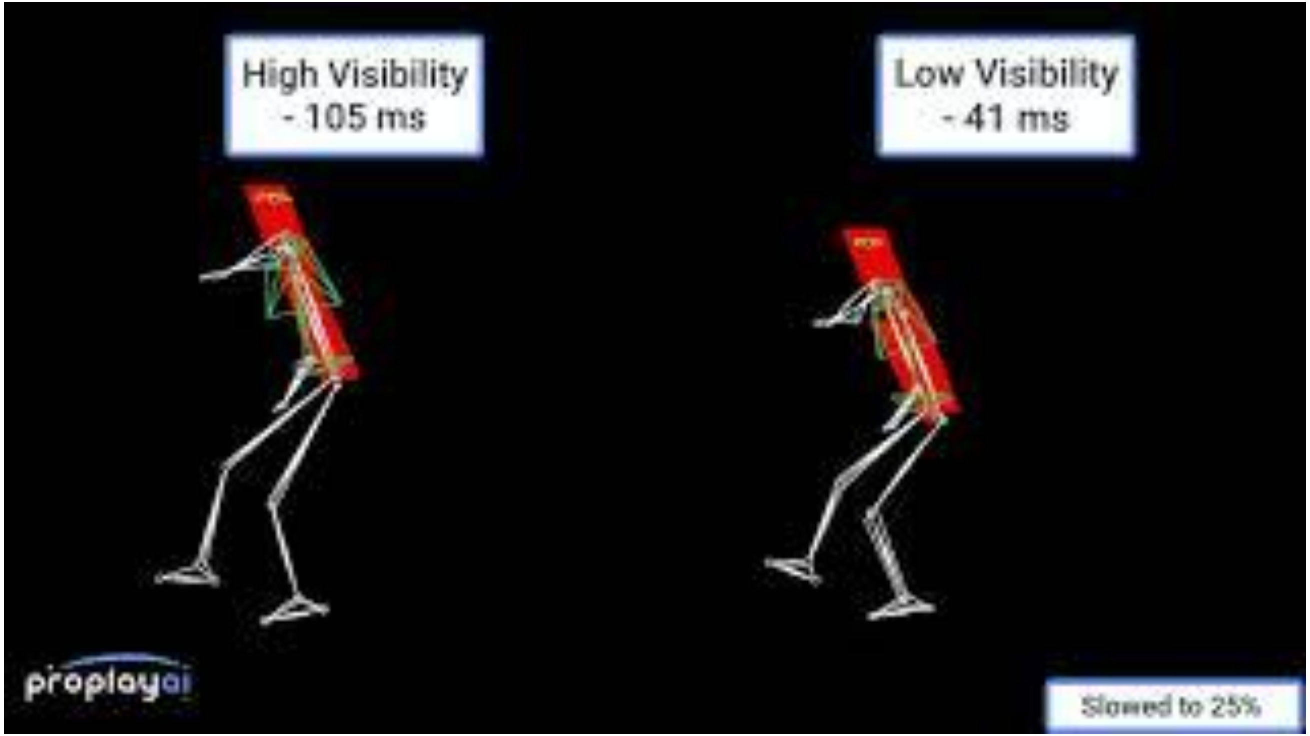
The Statcast Revolution

Big Data Baseball



The Statcast Revolution

Big Data Baseball

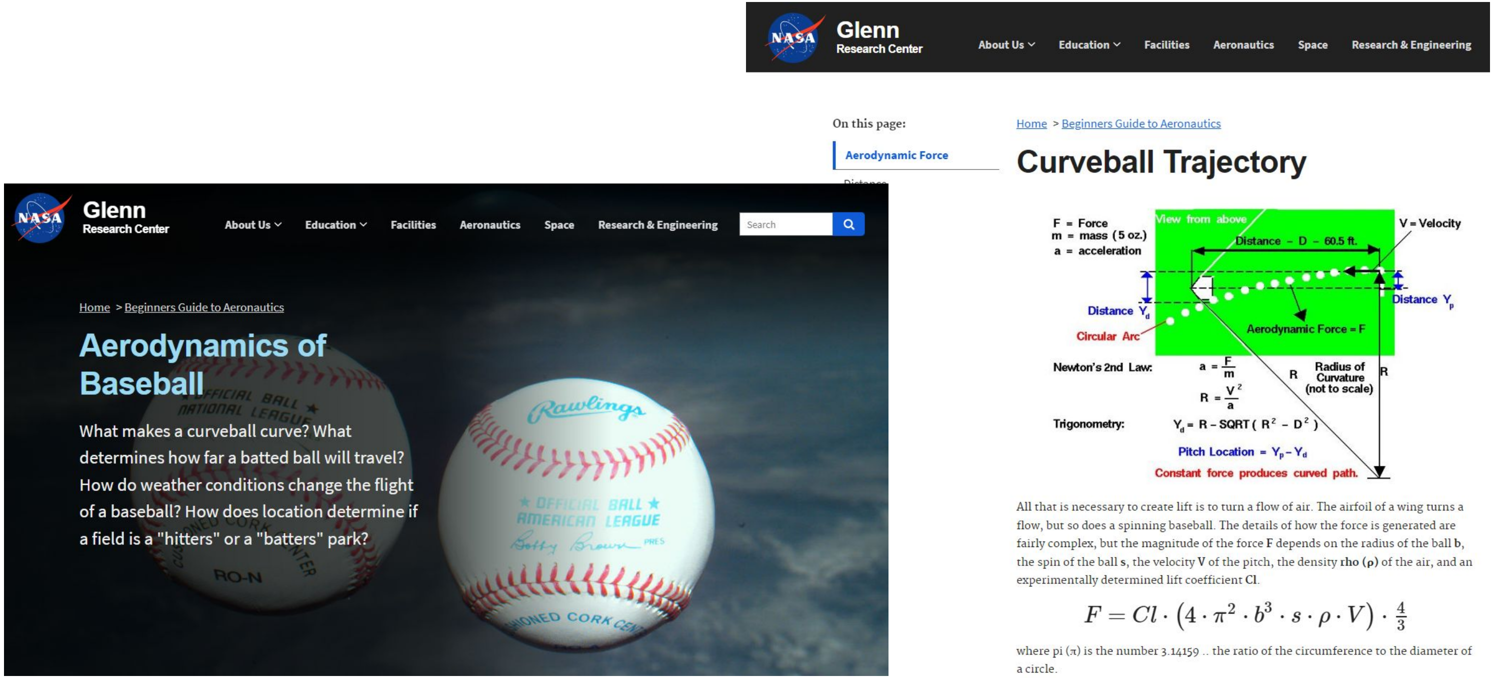


The Statcast Revolution

Big Data Baseball

The Statcast Revolution

Big Data Baseball



Home > Beginners Guide to Aeronautics

Aerodynamics of Baseball

What makes a curveball curve? What determines how far a batted ball will travel? How do weather conditions change the flight of a baseball? How does location determine if a field is a "hitters" or a "batters" park?

On this page:

Aerodynamic Force

Home > Beginners Guide to Aeronautics

Curveball Trajectory

All that is necessary to create lift is to turn a flow of air. The airfoil of a wing turns a flow, but so does a spinning baseball. The details of how the force is generated are fairly complex, but the magnitude of the force F depends on the radius of the ball **b**, the spin of the ball **s**, the velocity **V** of the pitch, the density rho (ρ) of the air, and an experimentally determined lift coefficient **Cl**.

$$F = Cl \cdot \left(4 \cdot \pi^2 \cdot b^3 \cdot s \cdot \rho \cdot V\right) \cdot \frac{4}{3}$$

where pi (π) is the number 3.14159 .. the ratio of the circumference to the diameter of a circle.

The Statcast Revolution

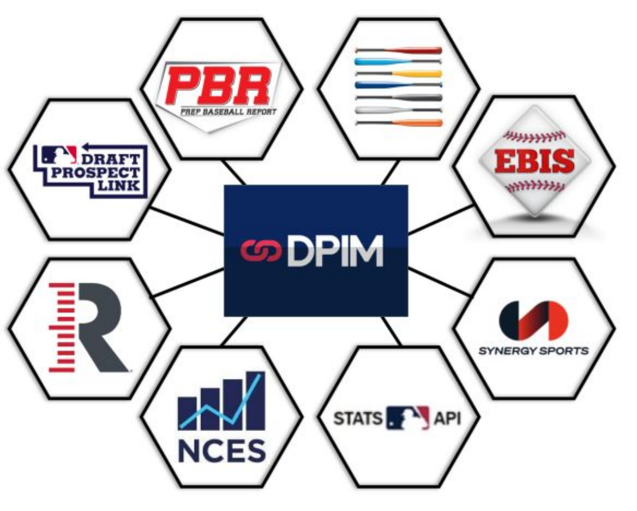
Big Data Baseball

CATAPULT

PROPLAYAI
LEARN TO WIN

Diamond KINETICS

TRACKMAN



simi
reality motion systems

ZELUS

BLAST

MAXPREPS

POLAR

SPARTASCIENCE

QUALISYS

KMOTION

HAWK-EYE
INNOVATIONS

PERFECT GAME

The Statcast Revolution

Big Data Baseball

Group A	Group B	Group C	Group D
ITALY	SPAIN	ISRAEL	THE NETHERLANDS
AUSTRIA	CROATIA	GREAT BRITAIN	SWEDEN
BELGIUM	GERMANY	FRANCE	CZECH REPUBLIC
QUALIFIER UTENA	QUALIFIER BELGRADO	QUALIFIER MOSCA	QUALIFIER BRATISLAVA

The Winning Formula: Upgrading With Airflow

“It’s math-y, but there’s still the whole arts-and-science debate [with big data]. I’d argue there is an art to that sort of stuff.”

Mike Fitzgerald
(VP of R&D for the Diamondbacks)

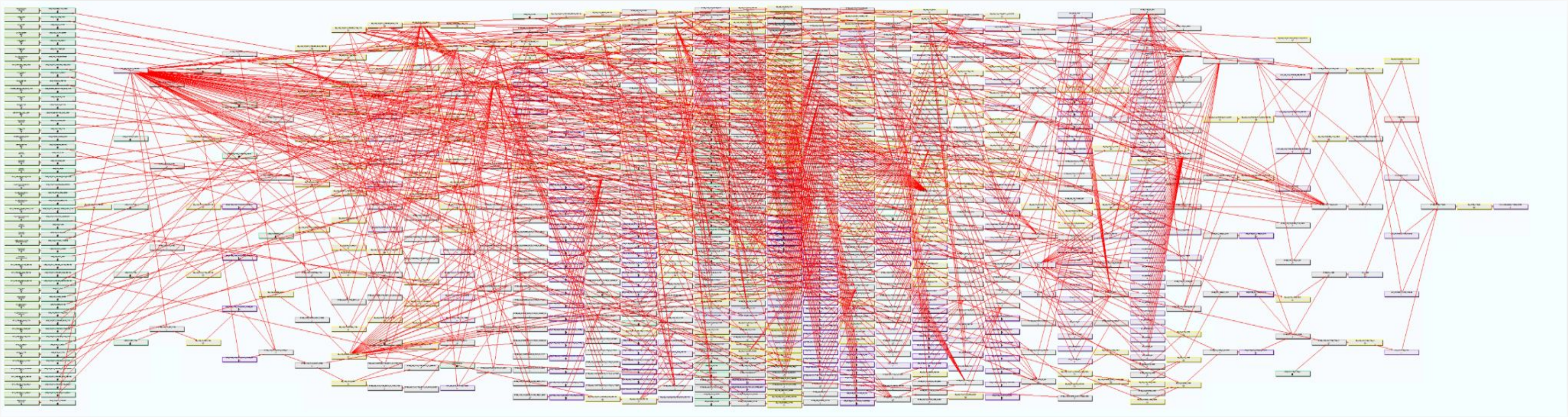
The Old Way

The Winning Formula

Our first data solution failed to adapt as we started ingesting big data.

Problems:

- Cron Scheduled
- Hard to troubleshoot
- Tough to explain



The Potential

The Winning Formula

- Consumers want data fast
- Individualized requirements
- Global consumers and schedules

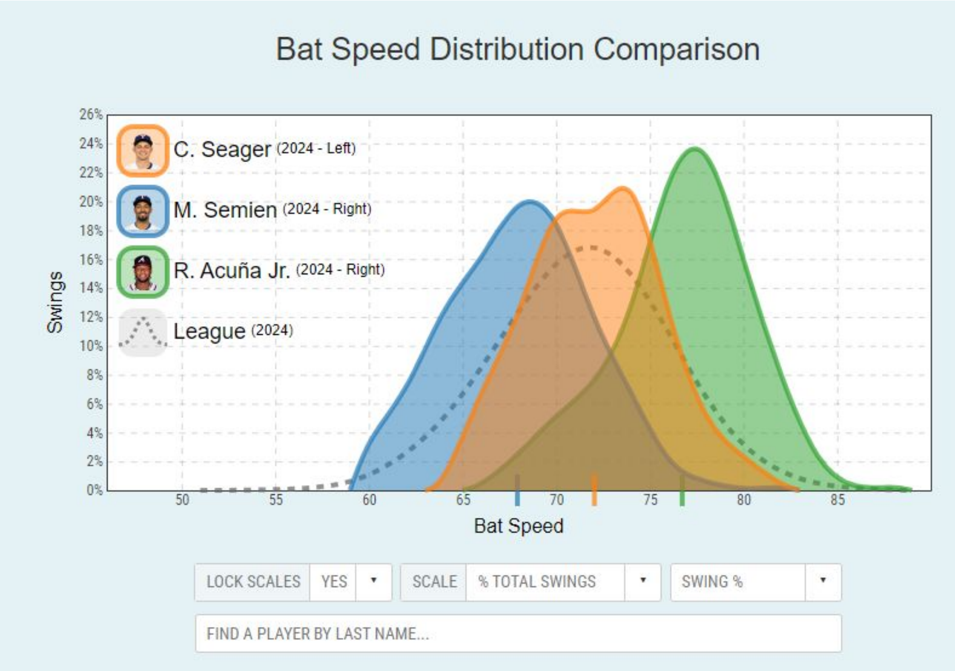
New Data Source Ingestion

The Winning Formula

Statcast **bat tracking** is available beginning with the 2024 season.

Since different parts of a bat can move at different speeds, an individual swing's speed is measured at the point six inches from the head of the bat, what is popularly called ***"the sweet-spot."***

Swing length tracks the sum distance traveled by the head of the bat in XYZ space from the start of data until contact point.

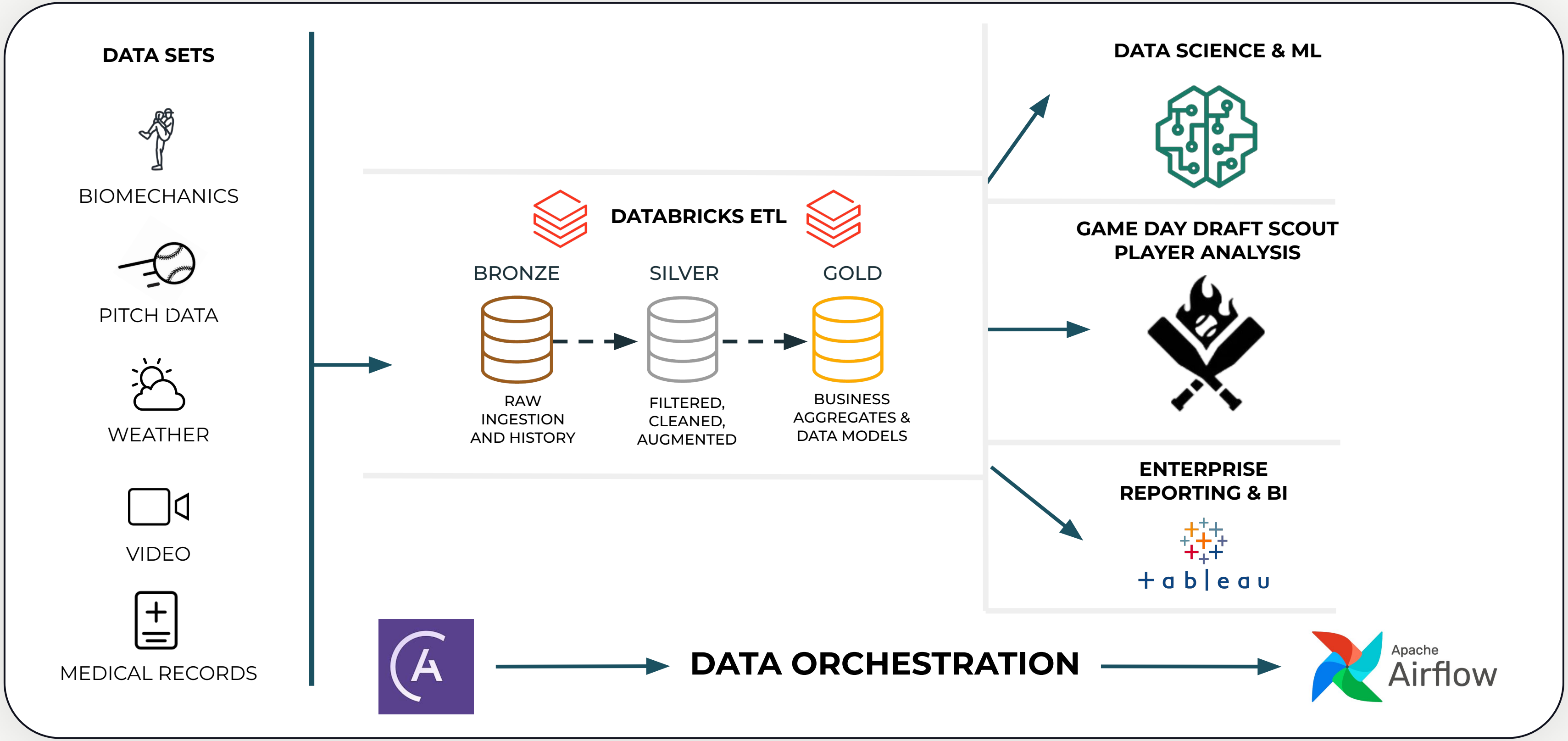


LOCK SCALES YES ▾ SCALE % TOTAL SWINGS ▾ SWING % ▾

FIND A PLAYER BY LAST NAME...

A Unified Data Platform

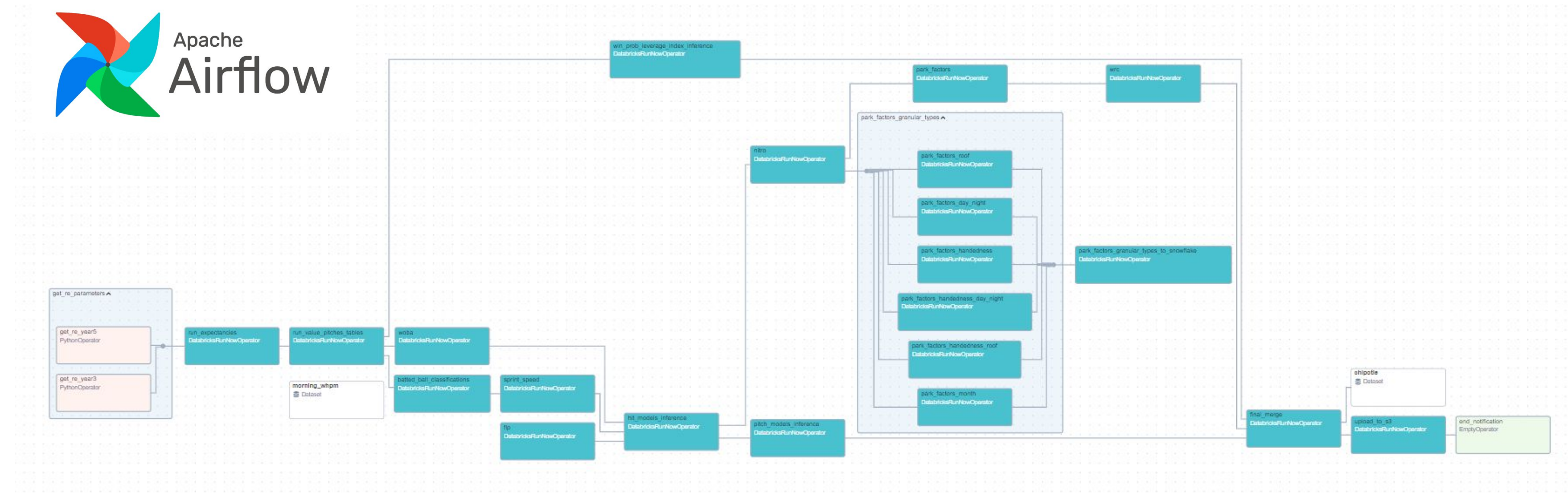
The Winning Formula



Upgrade With Airflow

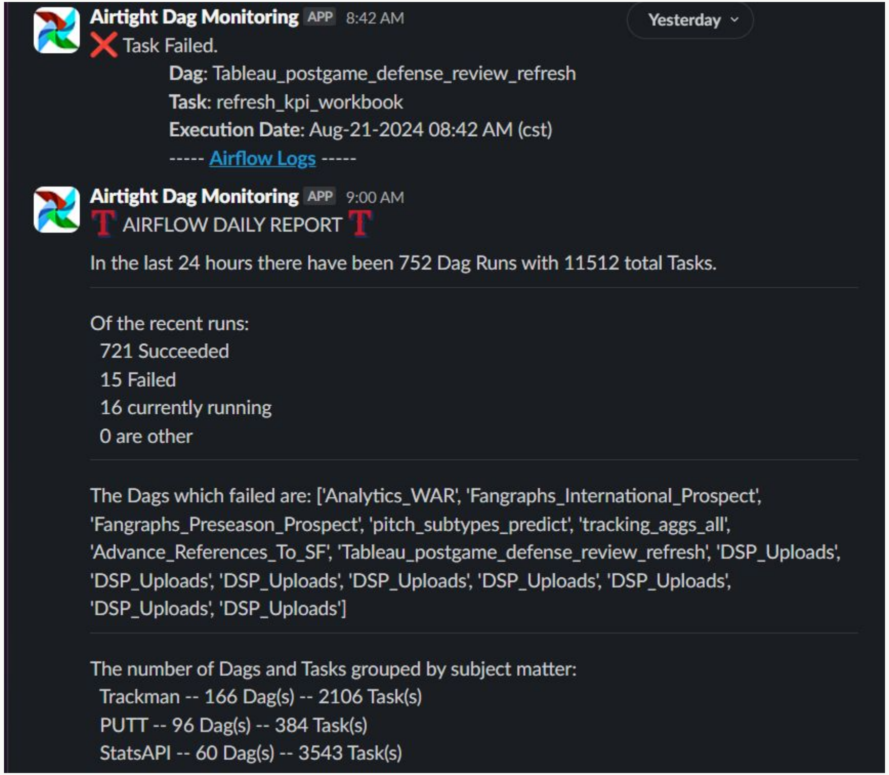
The Winning Formula

AIRFLOW SUMMIT 2024
10 YEARS OF AIRFLOW
SF·USA



Custom Monitoring

The Winning Formula



Airtight Dag Monitoring APP 8:42 AM

❌ Task Failed.

Dag: Tableau_postgame_defense_review_refresh
Task: refresh_kpi_workbook
Execution Date: Aug-21-2024 08:42 AM (cst)
----- Airflow Logs -----

Airtight Dag Monitoring APP 9:00 AM

T AIRFLOW DAILY REPORT T

In the last 24 hours there have been 752 Dag Runs with 11512 total Tasks.

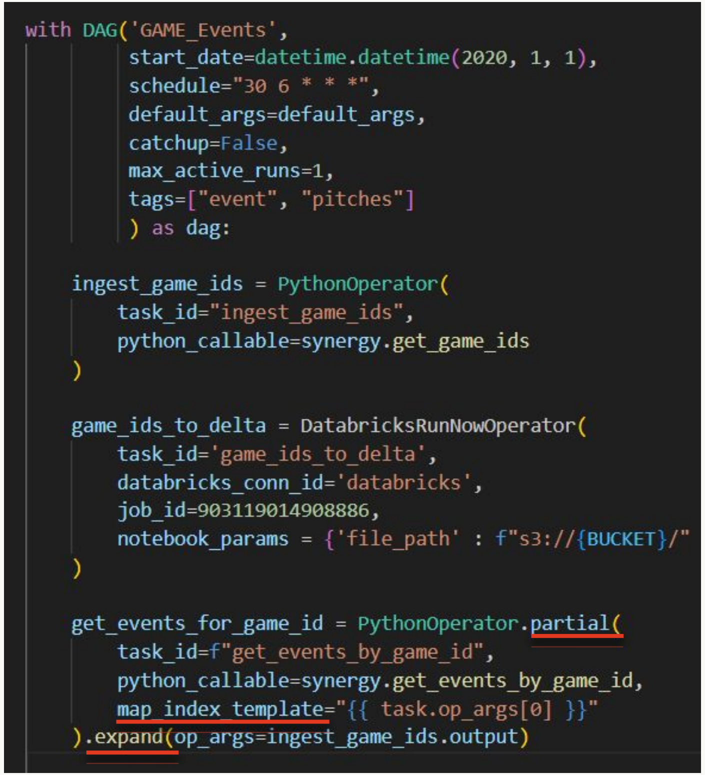
Of the recent runs:
721 Succeeded
15 Failed
16 currently running
0 are other

The Dags which failed are: ['Analytics_WAR', 'Fangraphs_International_Prospect', 'Fangraphs_Preseason_Prospect', 'pitch_subtypes_predict', 'tracking_aggs_all', 'Advance_References_To_SF', 'Tableau_postgame_defense_review_refresh', 'DSP_Uploads', 'DSP_Uploads', 'DSP_Uploads', 'DSP_Uploads', 'DSP_Uploads', 'DSP_Uploads', 'DSP_Uploads', 'DSP_Uploads']

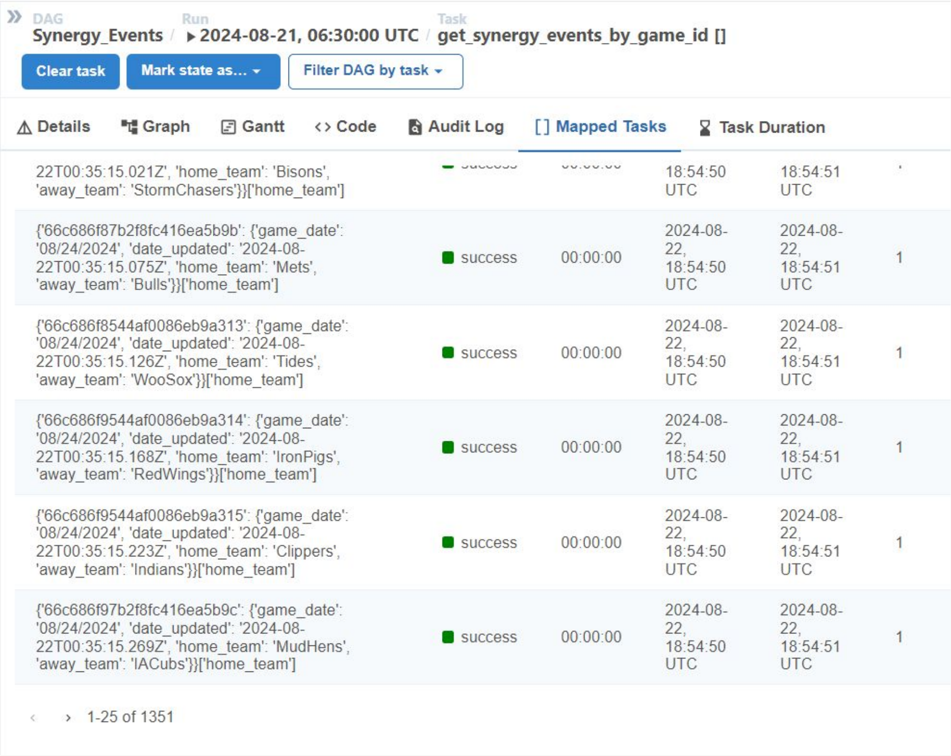
The number of Dags and Tasks grouped by subject matter:
Trackman -- 166 Dag(s) -- 2106 Task(s)
PUTT -- 96 Dag(s) -- 384 Task(s)
StatsAPI -- 60 Dag(s) -- 3543 Task(s)

Named Dynamic Tasks v2.9

The Winning Formula



```
with DAG('GAME_Events',
        start_date=datetime.datetime(2020, 1, 1),
        schedule="30 6 * * *",
        default_args=default_args,
        catchup=False,
        max_active_runs=1,
        tags=["event", "pitches"]
        ) as dag:

    ingest_game_ids = PythonOperator(
        task_id="ingest_game_ids",
        python_callable=synergy.get_game_ids
    )

    game_ids_to_delta = DatabricksRunNowOperator(
        task_id='game_ids_to_delta',
        databricks_conn_id='databricks',
        job_id=903119014908886,
        notebook_params = {'file_path' : f"s3://{BUCKET}/"
    )

    get_events_for_game_id = PythonOperator.partial(
        task_id=f"get_events_by_game_id",
        python_callable=synergy.get_events_by_game_id,
        map_index_template="{{ task.op_args[0] }}"
    ).expand(op_args=ingest_game_ids.output)
```

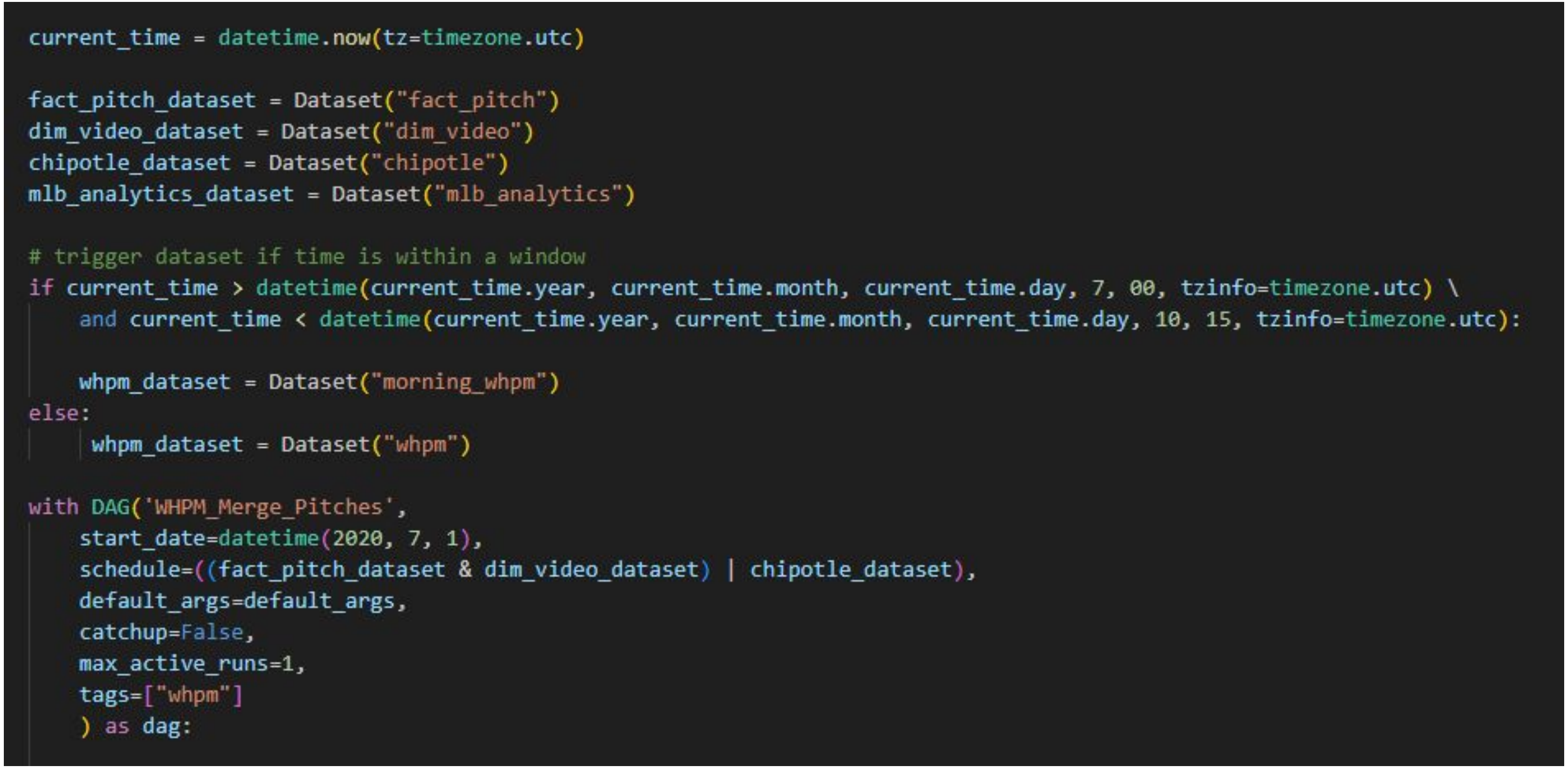


22T00:35:15.021Z', 'home_team': 'Bisons', 'away_team': 'StormChasers'}}['home_team']	success	00:00:00	18:54:50 UTC	18:54:51 UTC	1
{'66c686f87b2f8fc416ea5b9b': {'game_date': '08/24/2024', 'date_updated': '2024-08-22T00:35:15.075Z', 'home_team': 'Mets', 'away_team': 'Bulls'}}['home_team']	🟩 success	00:00:00	2024-08-22, 18:54:50 UTC	2024-08-22, 18:54:51 UTC	1
{'66c686f8544af0086eb9a313': {'game_date': '08/24/2024', 'date_updated': '2024-08-22T00:35:15.126Z', 'home_team': 'Tides', 'away_team': 'WooSox'}}['home_team']	🟩 success	00:00:00	2024-08-22, 18:54:50 UTC	2024-08-22, 18:54:51 UTC	1
{'66c686f9544af0086eb9a314': {'game_date': '08/24/2024', 'date_updated': '2024-08-22T00:35:15.168Z', 'home_team': 'IronPigs', 'away_team': 'RedWings'}}['home_team']	🟩 success	00:00:00	2024-08-22, 18:54:50 UTC	2024-08-22, 18:54:51 UTC	1
{'66c686f9544af0086eb9a315': {'game_date': '08/24/2024', 'date_updated': '2024-08-22T00:35:15.223Z', 'home_team': 'Clippers', 'away_team': 'Indians'}}['home_team']	🟩 success	00:00:00	2024-08-22, 18:54:50 UTC	2024-08-22, 18:54:51 UTC	1
{'66c686f97b2f8fc416ea5b9c': {'game_date': '08/24/2024', 'date_updated': '2024-08-22T00:35:15.269Z', 'home_team': 'MudHens', 'away_team': 'IACubs'}}['home_team']	🟩 success	00:00:00	2024-08-22, 18:54:50 UTC	2024-08-22, 18:54:51 UTC	1

‹ › 1-25 of 1351

Data Driven Scheduling v2.9

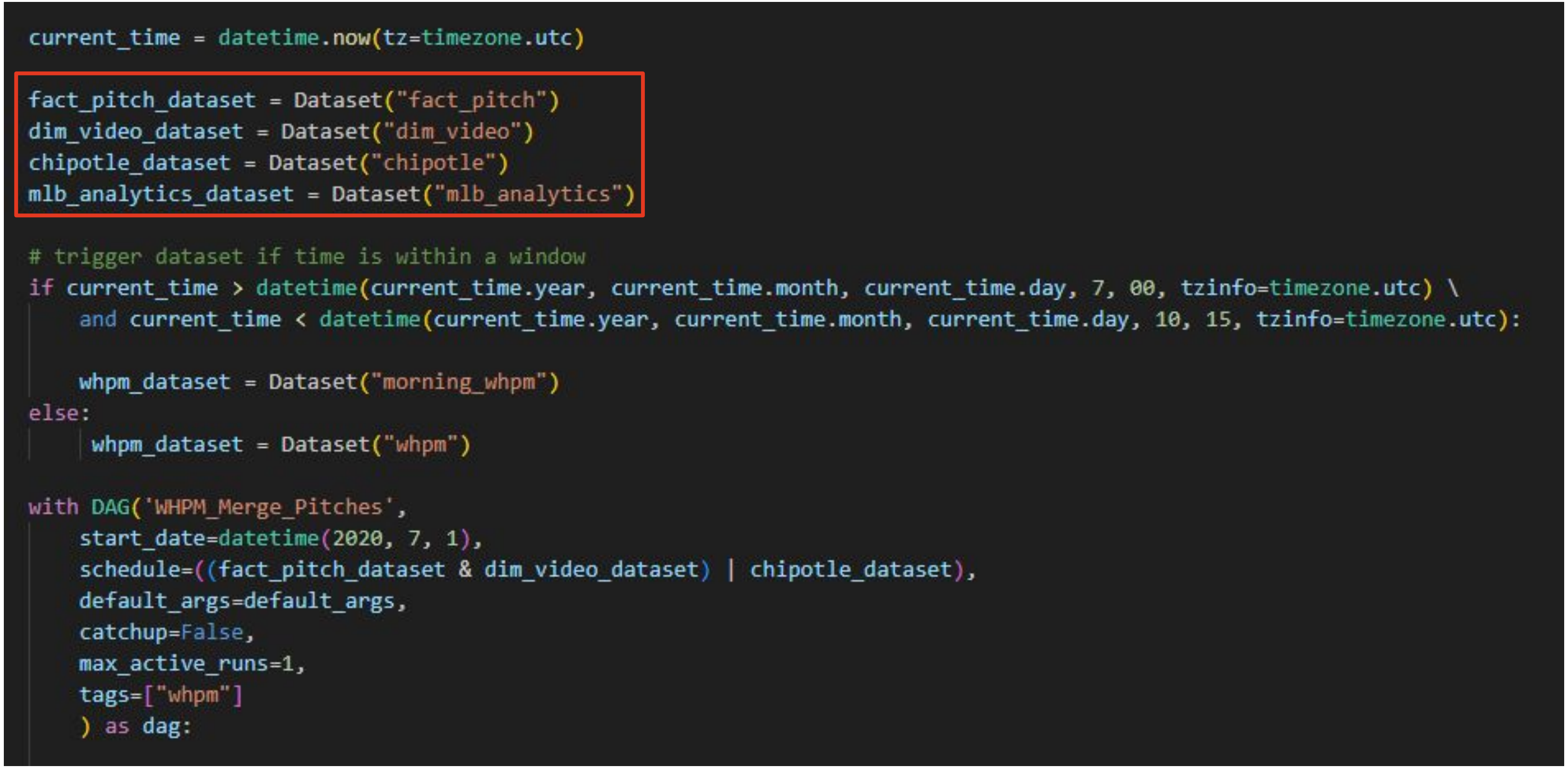
The Winning Formula



```python
current_time = datetime.now(tz=timezone.utc)

fact_pitch_dataset = Dataset("fact_pitch")
dim_video_dataset = Dataset("dim_video")
chipotle_dataset = Dataset("chipotle")
mlb_analytics_dataset = Dataset("mlb_analytics")

# trigger dataset if time is within a window
if current_time > datetime(current_time.year, current_time.month, current_time.day, 7, 00, tzinfo=timezone.utc) \
    and current_time < datetime(current_time.year, current_time.month, current_time.day, 10, 15, tzinfo=timezone.utc):

    whpm_dataset = Dataset("morning_whpm")
else:
     whpm_dataset = Dataset("whpm")

with DAG('WHPM_Merge_Pitches',
    start_date=datetime(2020, 7, 1),
    schedule=((fact_pitch_dataset & dim_video_dataset) | chipotle_dataset),
    default_args=default_args,
    catchup=False,
    max_active_runs=1,
    tags=["whpm"]
    ) as dag:
```

Data Driven Scheduling v2.9

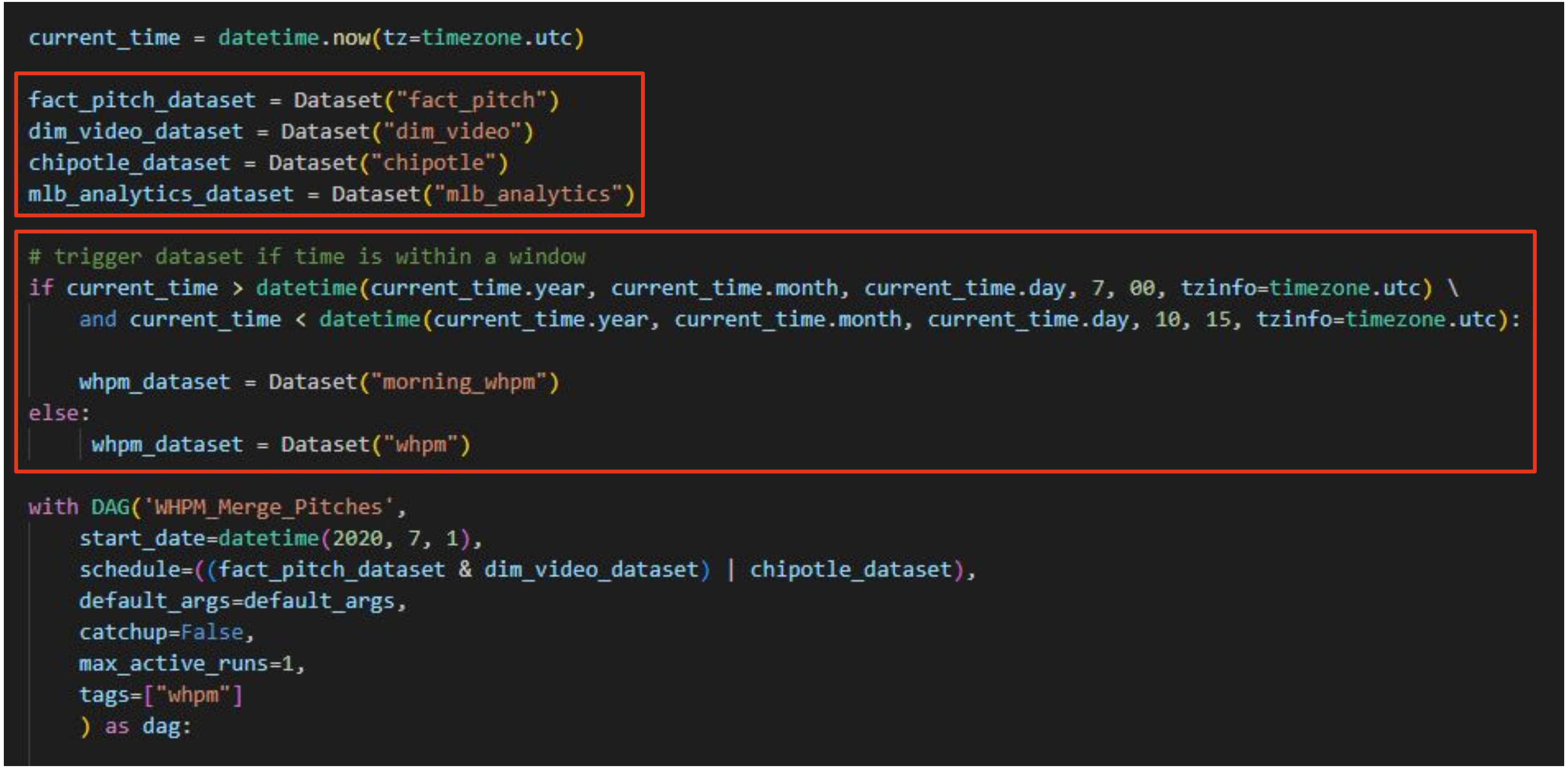
The Winning Formula



```python
current_time = datetime.now(tz=timezone.utc)

fact_pitch_dataset = Dataset("fact_pitch")
dim_video_dataset = Dataset("dim_video")
chipotle_dataset = Dataset("chipotle")
mlb_analytics_dataset = Dataset("mlb_analytics")

# trigger dataset if time is within a window
if current_time > datetime(current_time.year, current_time.month, current_time.day, 7, 00, tzinfo=timezone.utc) \
    and current_time < datetime(current_time.year, current_time.month, current_time.day, 10, 15, tzinfo=timezone.utc):

    whpm_dataset = Dataset("morning_whpm")
else:
    whpm_dataset = Dataset("whpm")

with DAG('WHPM_Merge_Pitches',
    start_date=datetime(2020, 7, 1),
    schedule=((fact_pitch_dataset & dim_video_dataset) | chipotle_dataset),
    default_args=default_args,
    catchup=False,
    max_active_runs=1,
    tags=["whpm"]
    ) as dag:
```

Data Driven Scheduling v2.9

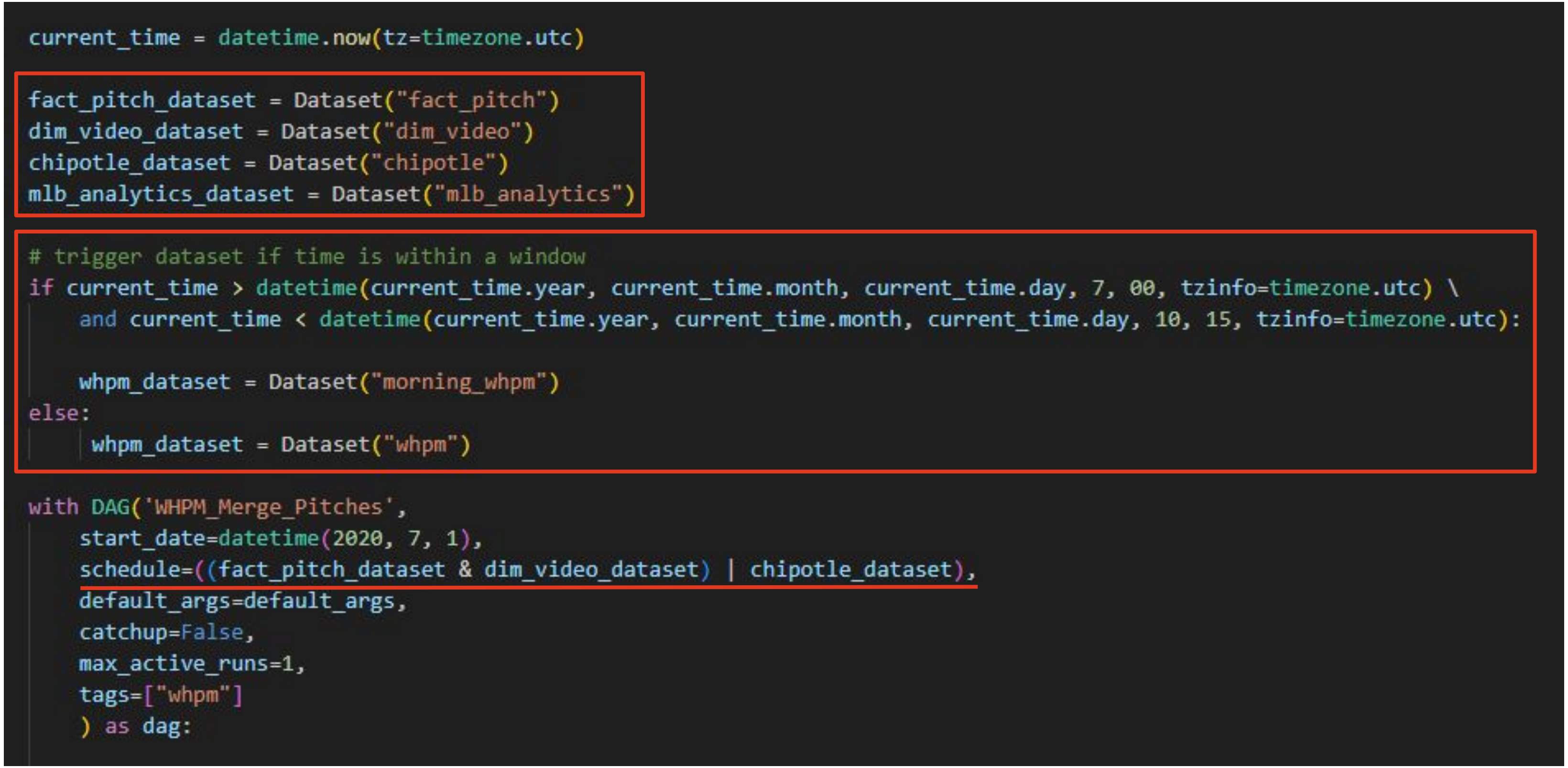
The Winning Formula



```
current_time = datetime.now(tz=timezone.utc)

fact_pitch_dataset = Dataset("fact_pitch")
dim_video_dataset = Dataset("dim_video")
chipotle_dataset = Dataset("chipotle")
mlb_analytics_dataset = Dataset("mlb_analytics")

# trigger dataset if time is within a window
if current_time > datetime(current_time.year, current_time.month, current_time.day, 7, 00, tzinfo=timezone.utc) \
    and current_time < datetime(current_time.year, current_time.month, current_time.day, 10, 15, tzinfo=timezone.utc):

    whpm_dataset = Dataset("morning_whpm")
else:
     whpm_dataset = Dataset("whpm")

with DAG('WHPM_Merge_Pitches',
    start_date=datetime(2020, 7, 1),
    schedule=((fact_pitch_dataset & dim_video_dataset) | chipotle_dataset),
    default_args=default_args,
    catchup=False,
    max_active_runs=1,
    tags=["whpm"]
    ) as dag:
```

Data Driven Scheduling v2.9

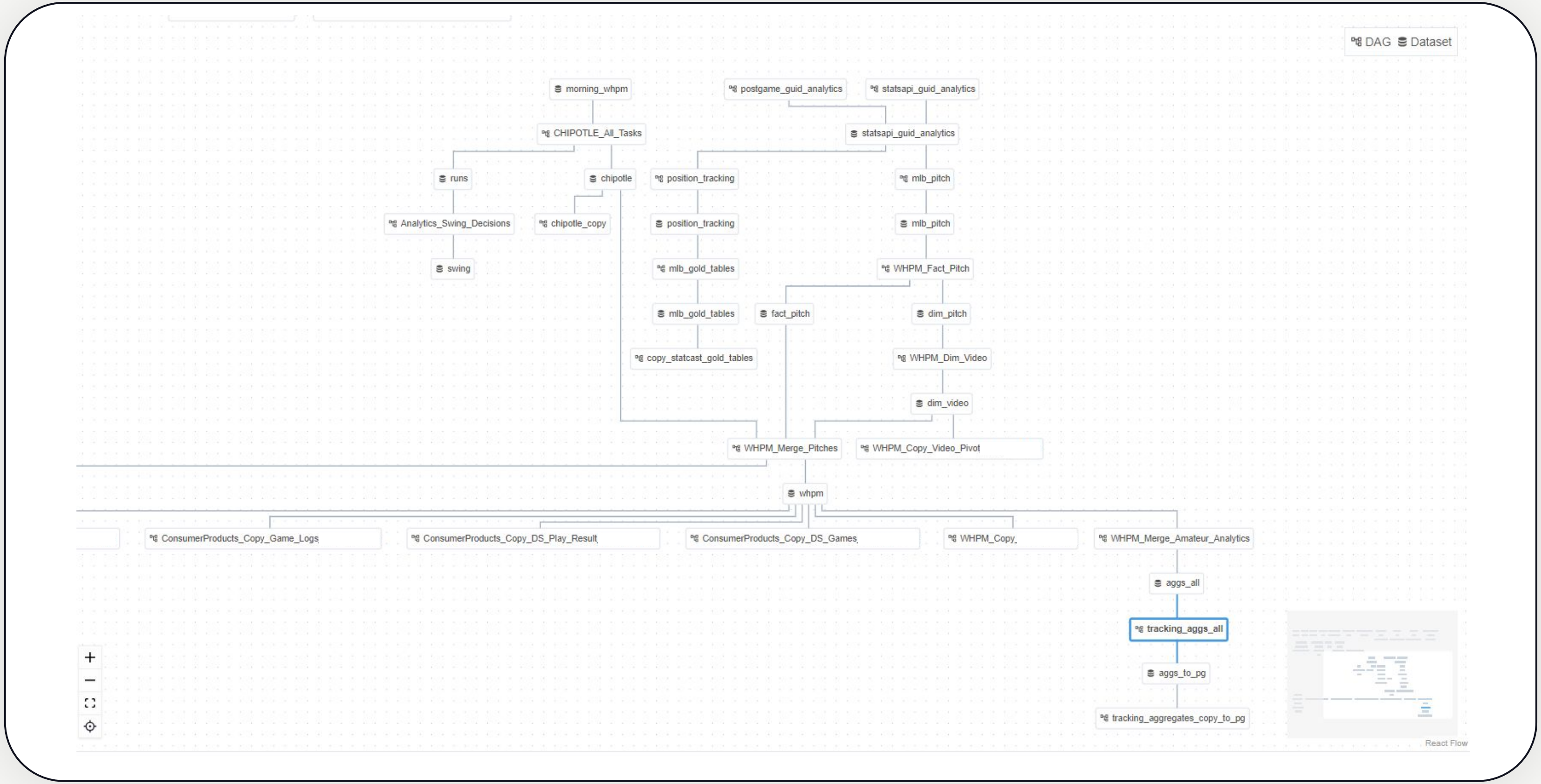
The Winning Formula



```
current_time = datetime.now(tz=timezone.utc)

fact_pitch_dataset = Dataset("fact_pitch")
dim_video_dataset = Dataset("dim_video")
chipotle_dataset = Dataset("chipotle")
mlb_analytics_dataset = Dataset("mlb_analytics")

# trigger dataset if time is within a window
if current_time > datetime(current_time.year, current_time.month, current_time.day, 7, 00, tzinfo=timezone.utc) \
    and current_time < datetime(current_time.year, current_time.month, current_time.day, 10, 15, tzinfo=timezone.utc):

    whpm_dataset = Dataset("morning_whpm")
else:
     whpm_dataset = Dataset("whpm")

with DAG('WHPM_Merge_Pitches',
    start_date=datetime(2020, 7, 1),
    schedule=((fact_pitch_dataset & dim_video_dataset) | chipotle_dataset),
    default_args=default_args,
    catchup=False,
    max_active_runs=1,
    tags=["whpm"]
    ) as dag:
```

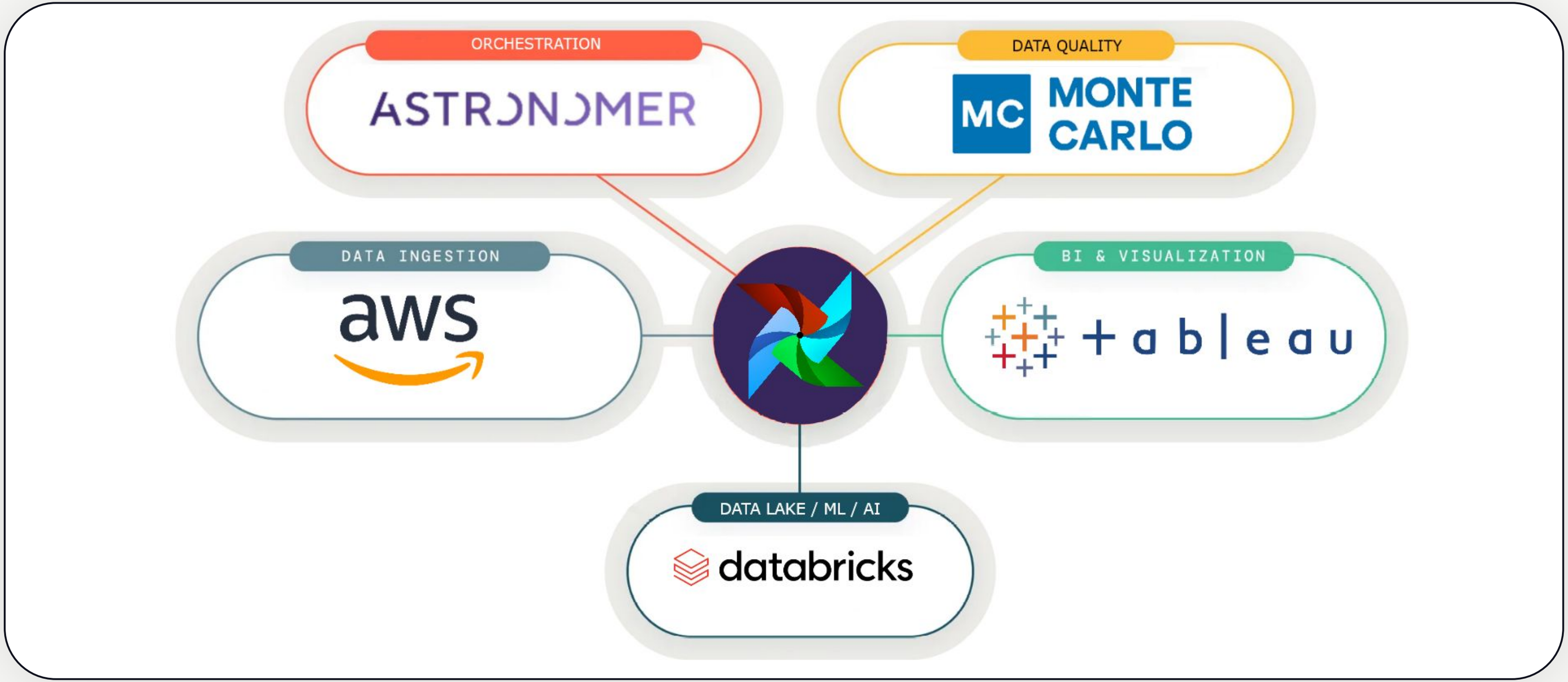
Data Driven Scheduling v2.9

The Winning Formula



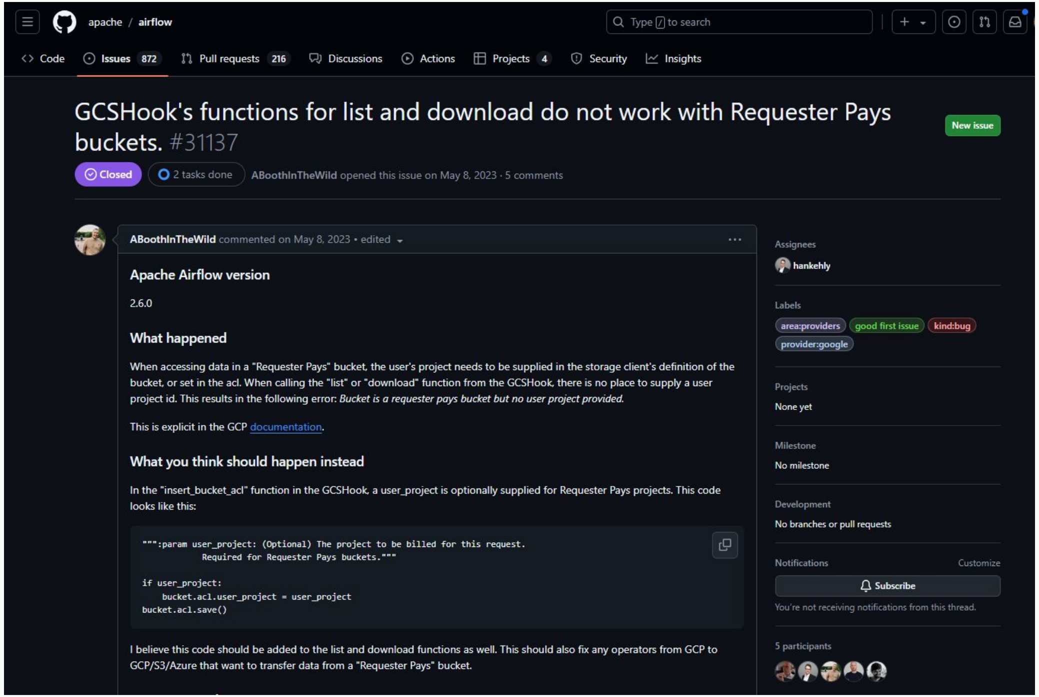
Partner Solutions

The Winning Formula



Open Source

The Winning Formula



GCSHook's functions for list and download do not work with Requester Pays buckets. #31137

Closed · 2 tasks done · ABoothInTheWild opened this issue on May 8, 2023 · 5 comments

ABoothInTheWild commented on May 8, 2023 • edited

Apache Airflow version

2.6.0

What happened

When accessing data in a "Requester Pays" bucket, the user's project needs to be supplied in the storage client's definition of the bucket, or set in the acl. When calling the "list" or "download" function from the GCSHook, there is no place to supply a user project id. This results in the following error: *Bucket is a requester pays bucket but no user project provided.*

This is explicit in the GCP documentation.

What you think should happen instead

In the "insert_bucket_acl" function in the GCSHook, a user_project is optionally supplied for Requester Pays projects. This code looks like this:

```
""":param user_project: (Optional) The project to be billed for this request.
            Required for Requester Pays buckets."""

if user_project:
    bucket.acl.user_project = user_project
bucket.acl.save()
```

I believe this code should be added to the list and download functions as well. This should also fix any operators from GCP to GCP/S3/Azure that want to transfer data from a "Requester Pays" bucket.

Assignees: hankehly

Labels: area:providers, good first issue, kind:bug, provider:google

Projects: None yet

Milestone: No milestone

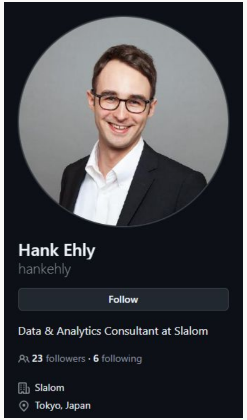
Development: No branches or pull requests

Notifications: Customize

Subscribe

You're not receiving notifications from this thread.

5 participants



Hank Ehly
hankehly

Data & Analytics Consultant at Slalom

23 followers · 6 following

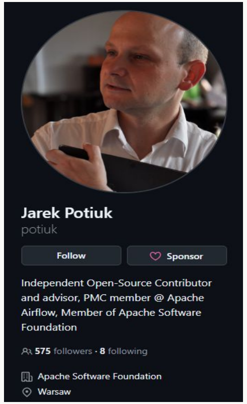
Slalom
Tokyo, Japan

Elad Kalif
eladkal

Apache-Airflow committer & PMC member. Data Engineer at Amazon

60 followers · 0 following

Amazon
Tel Aviv, Israel



Jarek Potiuk
potiuk

Independent Open-Source Contributor and advisor, PMC member @ Apache Airflow, Member of Apache Software Foundation

575 followers · 8 following

Apache Software Foundation
Warsaw

Pankaj Singh
pankajastro

Software Engineer

12 followers · 2 following

Astronomer
India

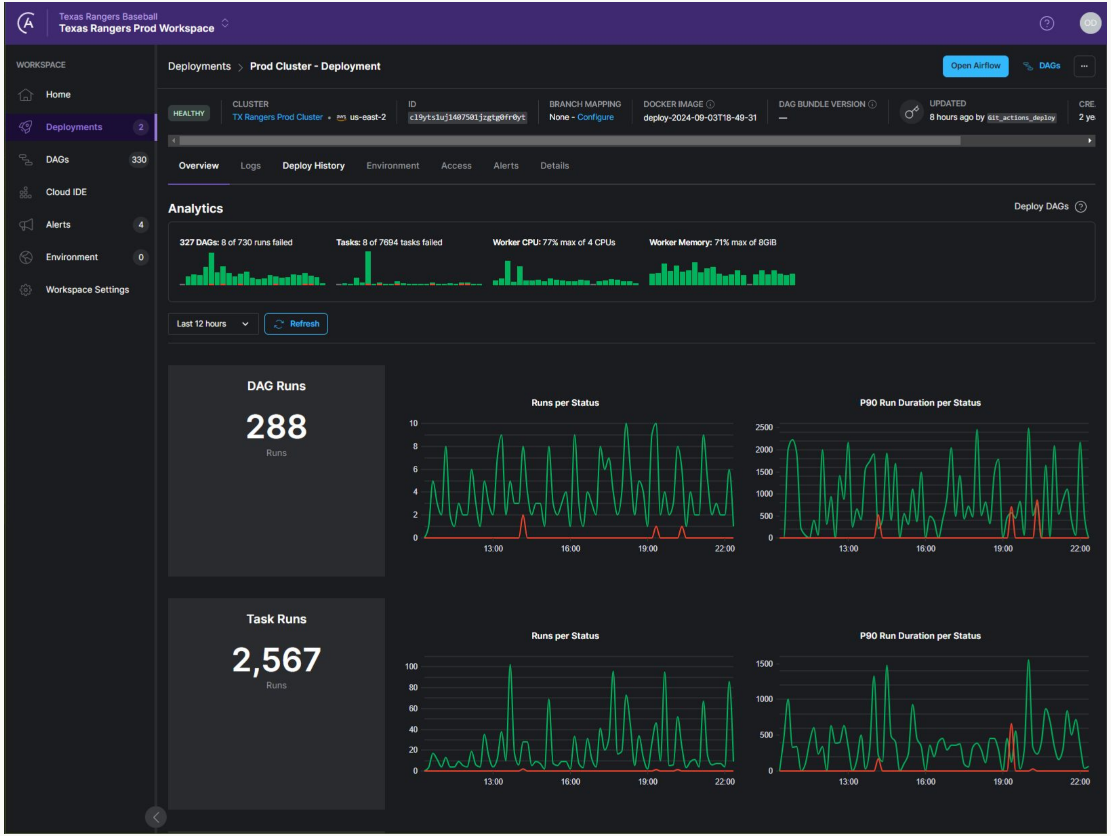
The Winning Formula: Astronomer Take the Wheel

"Half of it's art... it's creativity, and then half of it is just knowing the data you're working with and being able to manipulate it in the direction that will benefit the player."

Brian Bannister
(Director of Pitching,
Chicago White Sox)

Managed Airflow Environment

The Winning Formula



Texas Rangers Baseball
Texas Rangers Prod Workspace

Deployments > Prod Cluster - Deployment

HEALTHY

CLUSTER: TX Rangers Prod Cluster • aws us-east-2

ID: c19yts1uj1407501jzgtg0fr0yt

BRANCH MAPPING: None - Configure

DOCKER IMAGE: deploy-2024-09-03T18-49-31

DAG BUNDLE VERSION: —

UPDATED: 8 hours ago by Git_actions_deploy

CRE. 2 ye

Analytics

327 DAGs: 8 of 730 runs failed

Tasks: 8 of 7694 tasks failed

Worker CPU: 77% max of 4 CPUs

Worker Memory: 71% max of 8GiB

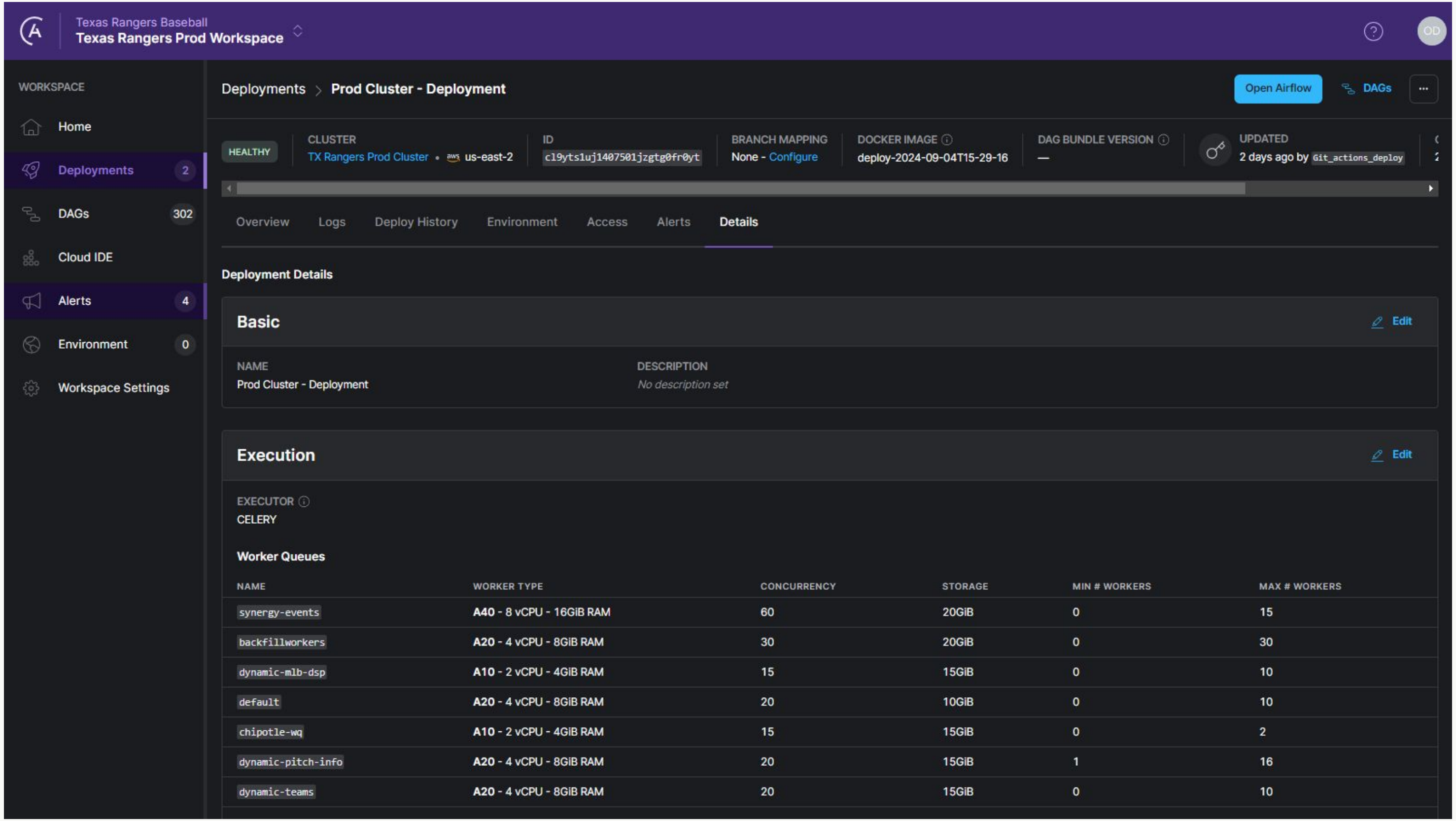
Last 12 hours

DAG Runs: 288 Runs

Task Runs: 2,567 Runs

Managed Airflow Environment

The Winning Formula



HEALTHY

CLUSTER: TX Rangers Prod Cluster • aws us-east-2

ID: c19yts1uj1407501jzgtg0fr0yt

BRANCH MAPPING: None - Configure

DOCKER IMAGE: deploy-2024-09-04T15-29-16

DAG BUNDLE VERSION: —

UPDATED: 2 days ago by Git_actions_deploy

Deployment Details

Basic

NAME: Prod Cluster - Deployment

DESCRIPTION: *No description set*

Execution

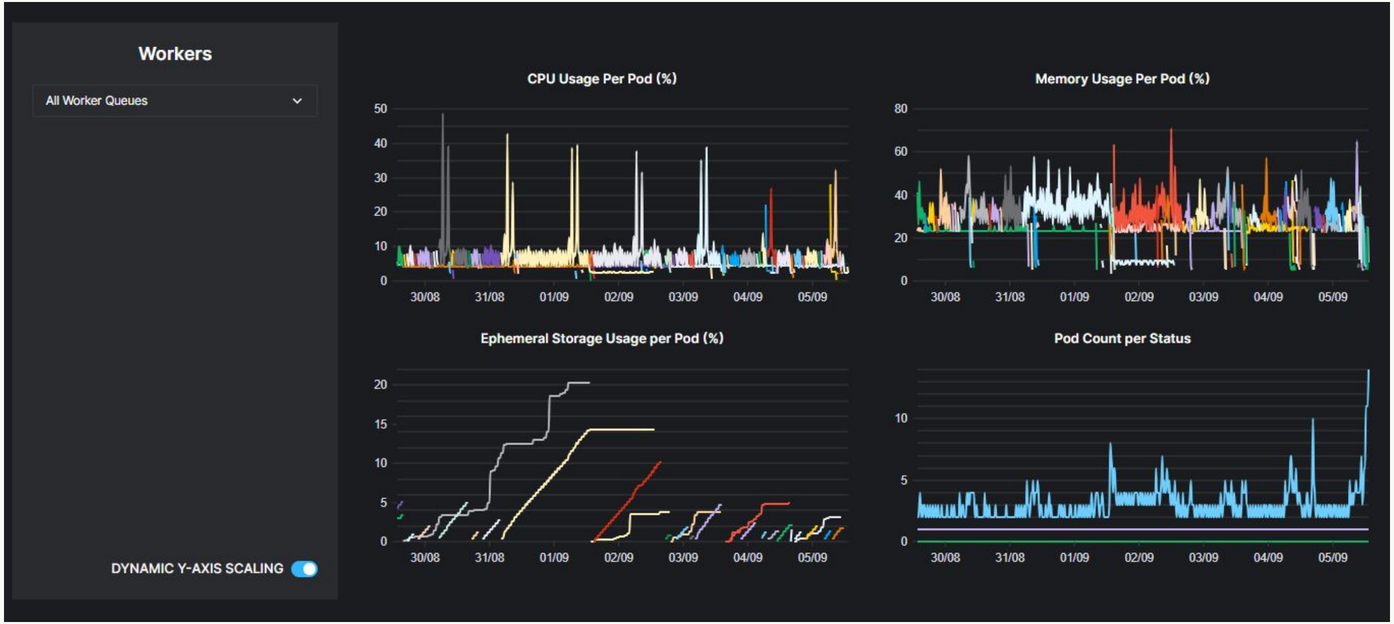
EXECUTOR: CELERY

Worker Queues

NAME	WORKER TYPE	CONCURRENCY	STORAGE	MIN # WORKERS	MAX # WORKERS
synergy-events	A40 - 8 vCPU - 16GiB RAM	60	20GiB	0	15
backfillworkers	A20 - 4 vCPU - 8GiB RAM	30	20GiB	0	30
dynamic-mlb-dsp	A10 - 2 vCPU - 4GiB RAM	15	15GiB	0	10
default	A20 - 4 vCPU - 8GiB RAM	20	10GiB	0	10
chipotle-wq	A10 - 2 vCPU - 4GiB RAM	15	15GiB	0	2
dynamic-pitch-info	A20 - 4 vCPU - 8GiB RAM	20	15GiB	1	16
dynamic-teams	A20 - 4 vCPU - 8GiB RAM	20	15GiB	0	10

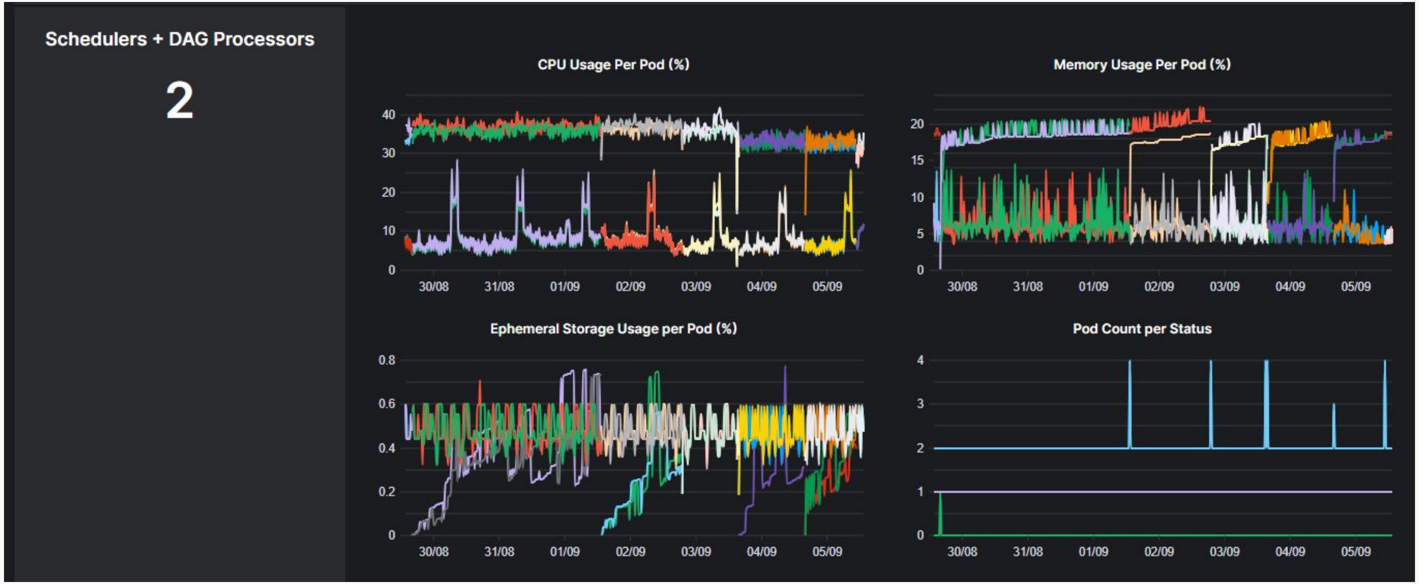
Managed Airflow Environment

The Winning Formula



Managed Airflow Environment

The Winning Formula

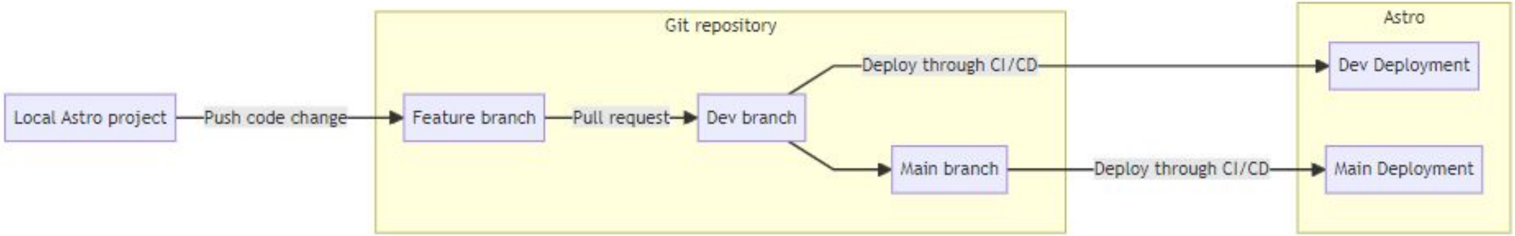


Managed Airflow Environment

The Winning Formula

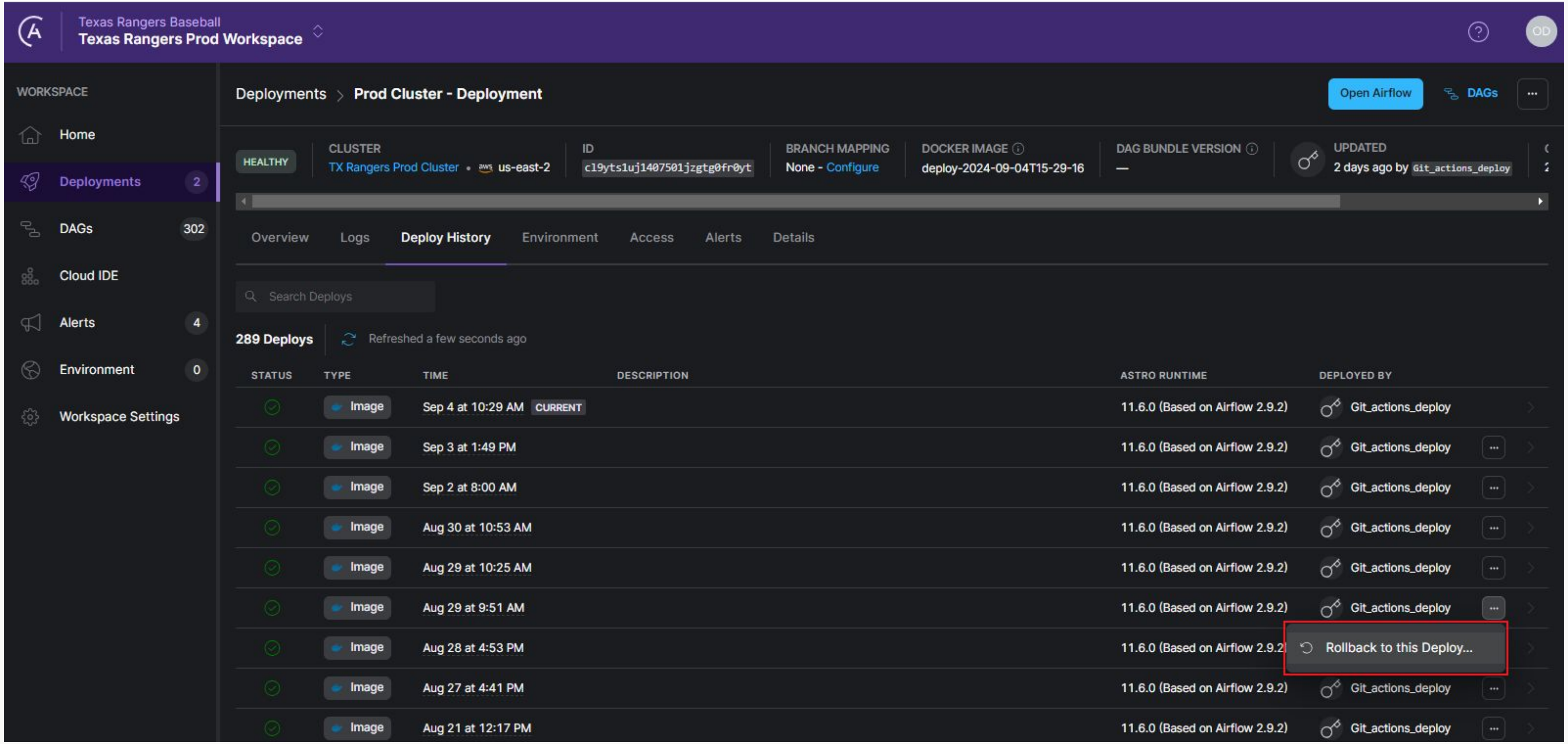
CI/CD with Astronomer–

- Guaranteed Code Review
- Automatic Tests and Deployment
- Easy Dev Promotion Process



The Statcast Revolution

The Winning Formula



Texas Rangers Baseball
Texas Rangers Prod Workspace

WORKSPACE

- Home
- Deployments 2
- DAGs 302
- Cloud IDE
- Alerts 4
- Environment 0
- Workspace Settings

Deployments > Prod Cluster - Deployment

Open Airflow | DAGs | ···

HEALTHY

CLUSTER	ID	BRANCH MAPPING	DOCKER IMAGE	DAG BUNDLE VERSION	UPDATED
TX Rangers Prod Cluster • aws us-east-2	cl9yts1uj1407501jzgtg0fr0yt	None - Configure	deploy-2024-09-04T15-29-16	—	2 days ago by Git_actions_deploy

Overview | Logs | **Deploy History** | Environment | Access | Alerts | Details

Search Deploys

289 Deploys Refreshed a few seconds ago

STATUS	TYPE	TIME	DESCRIPTION	ASTRO RUNTIME	DEPLOYED BY
✓	Image	Sep 4 at 10:29 AM CURRENT		11.6.0 (Based on Airflow 2.9.2)	Git_actions_deploy
✓	Image	Sep 3 at 1:49 PM		11.6.0 (Based on Airflow 2.9.2)	Git_actions_deploy ···
✓	Image	Sep 2 at 8:00 AM		11.6.0 (Based on Airflow 2.9.2)	Git_actions_deploy ···
✓	Image	Aug 30 at 10:53 AM		11.6.0 (Based on Airflow 2.9.2)	Git_actions_deploy ···
✓	Image	Aug 29 at 10:25 AM		11.6.0 (Based on Airflow 2.9.2)	Git_actions_deploy ···
✓	Image	Aug 29 at 9:51 AM		11.6.0 (Based on Airflow 2.9.2)	Git_actions_deploy ···
✓	Image	Aug 28 at 4:53 PM		11.6.0 (Based on Airflow 2.9.2)	Rollback to this Deploy...
✓	Image	Aug 27 at 4:41 PM		11.6.0 (Based on Airflow 2.9.2)	Git_actions_deploy ···
✓	Image	Aug 21 at 12:17 PM		11.6.0 (Based on Airflow 2.9.2)	Git_actions_deploy ···

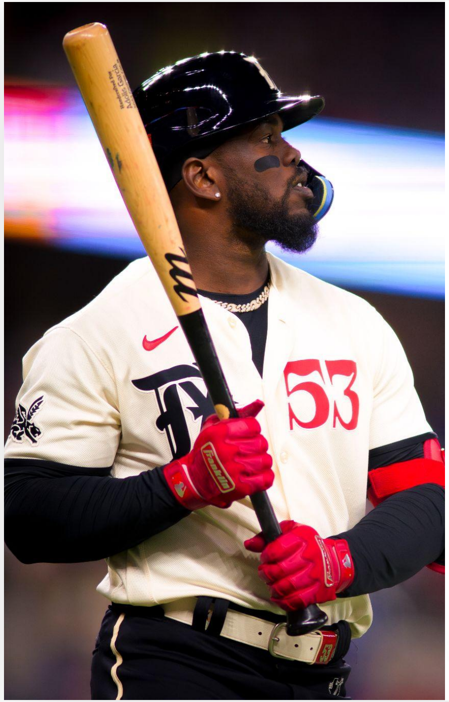
Putting It All Together: World Series Victory

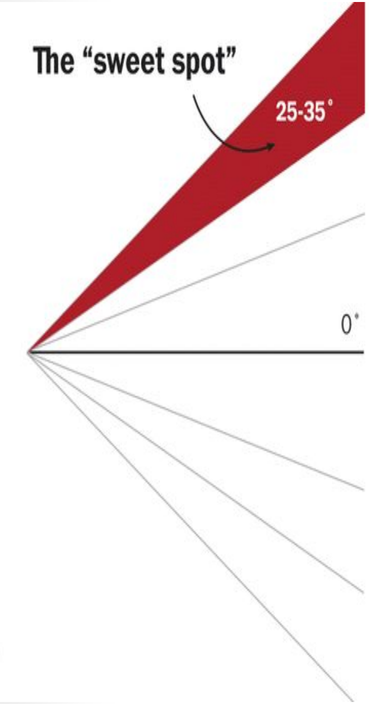
“If you can make yourself half a percent better… then that’s a win. It can be the difference between making the playoffs and not making the playoffs.”

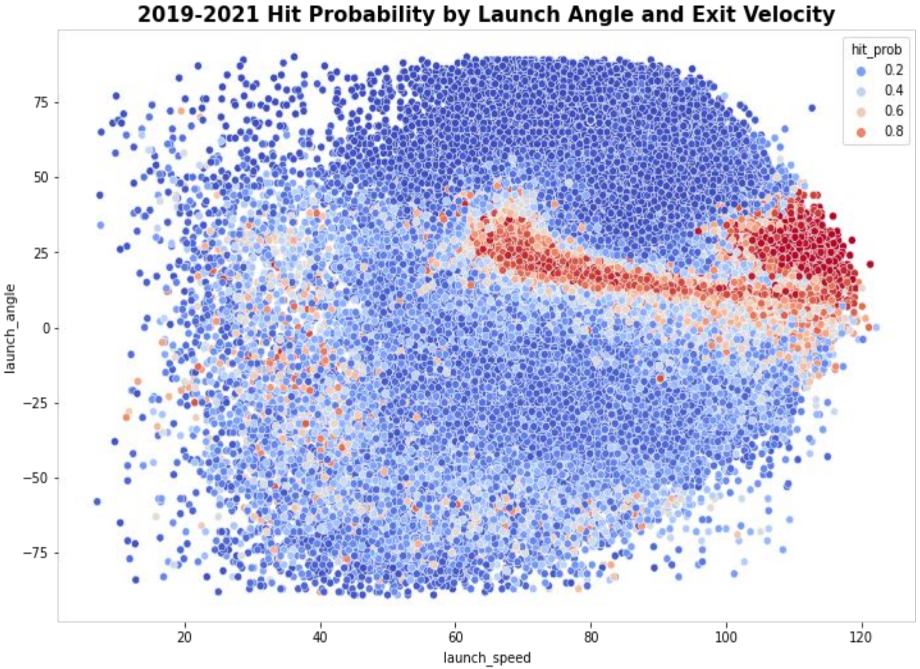
Ryan Murray
(Sr. Director of Baseball R&D,
Texas Rangers)

The Launch Angle Revolution

Putting It All Together

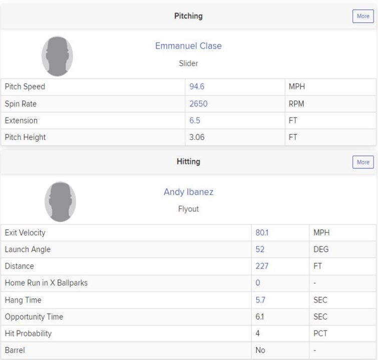






Reports and Visualizations

Putting It All Together



Pitching More

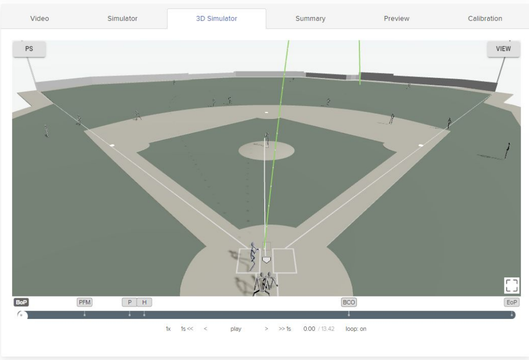
Emmanuel Clase
Slider

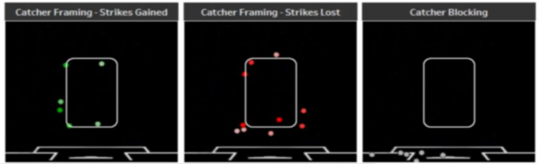
Pitch Speed	94.6	MPH
Spin Rate	2650	RPM
Extension	6.5	FT
Pitch Height	3.06	FT

Hitting More

Andy Ibanez
Flyout

Exit Velocity	80.1	MPH
Launch Angle	52	DEG
Distance	227	FT
Home Run in X Ballparks	0	-
Hang Time	5.7	SEC
Opportunity Time	6.1	SEC
Hit Probability	4	PCT
Barrel	No	-





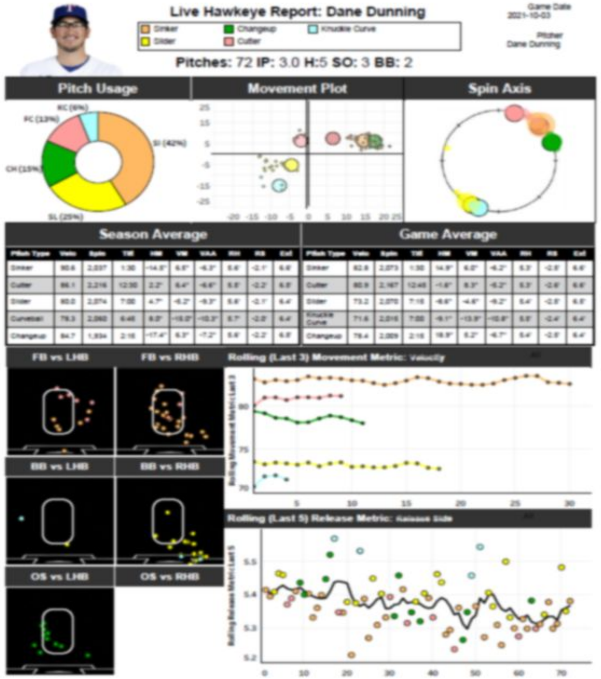
Inning	Outs	Pitcher	Batter	Count	Pitch Type	Strike Prob	Call Desc	
1	0	Dane Dunning	Jose Ramirez	0-0	Cutter	15.1%	Ball - Called	-.15
				1-0	Cutter	86.8%	Ball - Called	-.87
2	0	Dane Dunning	Bradley Zimmer	0-0	Changeup	40.0%	Ball - Called	-.40
	1	Dane Dunning	Oscar Mercado	2-2	Sinker	13.5%	Ball - Called	-.14
3	0	Dane Dunning	Amed Rosario	0-0	Slider	55.4%	Ball - Called	-.55
				1-0	Slider	19.4%	Ball - Called	-.19
				2-0	Sinker	10.8%	Ball - Called	-.11
				3-1	Cutter	42.2%	Ball - Called	-.42
			Jose Ramirez	0-0	Sinker	93.0%	Strike - Called	.07
	1	Dane Dunning	Yu Chang	2-2	Sinker	94.1%	Strike - Called	.06
	2	Dane Dunning	Austin Hedges	0-1	Sinker	9.0%	Ball - Called	-.09
			Bradley Zimmer	2-0	Sinker	45.0%	Strike - Called	.55
4	1	Mike Foltynewicz	Myles Straw	0-0	Fastball	6.5%	Ball - Called	-.06
				1-1	Curveball	92.2%	Strike - Called	.08
	2	Mike Foltynewicz	Amed Rosario	0-0	Fastball	59.0%	Strike - Called	.41
5	1	Mike Foltynewicz	Bradley Zimmer	0-0	Curveball	84.5%	Strike - Called	.15
			Franmil Reyes	3-0	Fastball	7.2%	Ball - Called	-.07
			Oscar Mercado	0-2	Curveball	36.1%	Ball - Called	-.36
6	1	Matt Bush	Myles Straw	1-2	Fastball	89.2%	Ball - Called	-.89
8	1	Jharel Cotton	Oscar Mercado	0-0	Fastball	52.6%	Strike - Called	.47
	2	Jharel Cotton	Austin Hedges	1-0	Fastball	9.4%	Ball - Called	-.09
9	0	Drew Anderson	Myles Straw	0-0	Slider	76.8%	Strike - Called	.23
	2	Drew Anderson	Jose Ramirez	2-0	Curveball	86.7%	Strike - Called	.13

Game Calling By Count

Count Desc	Count	FF	SI	FC	SL	CU	CH
1st Pitch	0-0	21%	21%	11%	21%	21%	5%
2nd Pitch	0-1	11%	17%		22%	17%	33%
	1-0	11%	50%	6%	11%	11%	11%
Ahead	0-2	20%	20%		30%	10%	20%
	1-2	13%	13%	6%	38%	13%	13%
	2-2	18%	27%		27%	18%	9%
Behind	2-0	11%	44%	11%	11%	22%	
	2-1		38%	13%	25%		25%
	3-0	67%	33%				
	3-1	25%	50%	25%			
Even	1-1	17%	17%		58%	8%	
Full	3-2	40%	20%		40%		

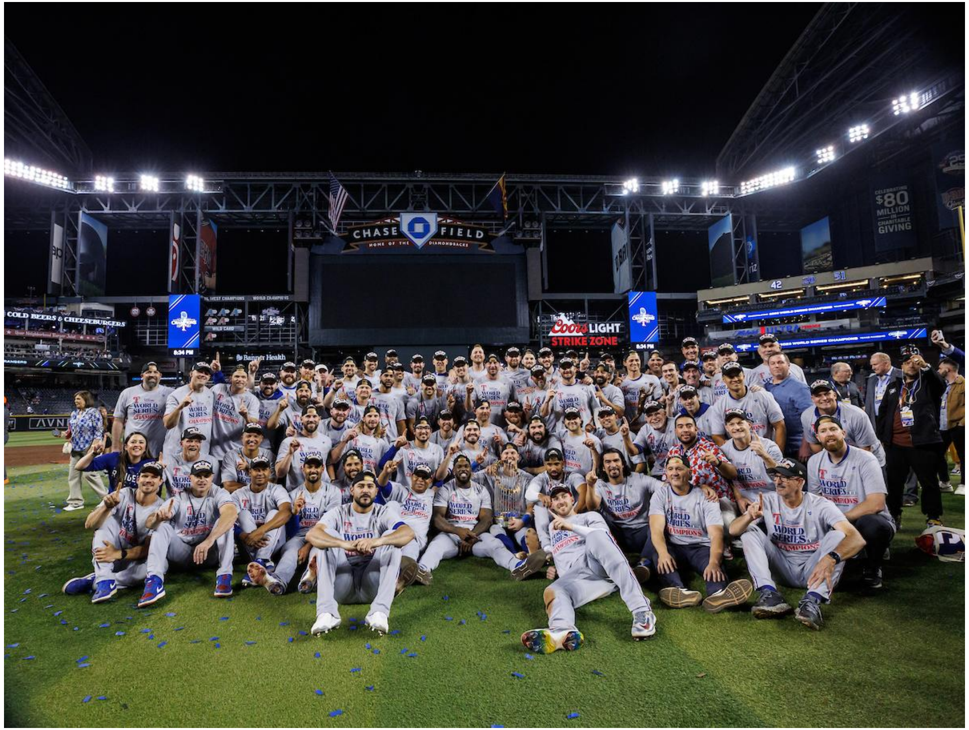
Game Calling By Prev Pitch

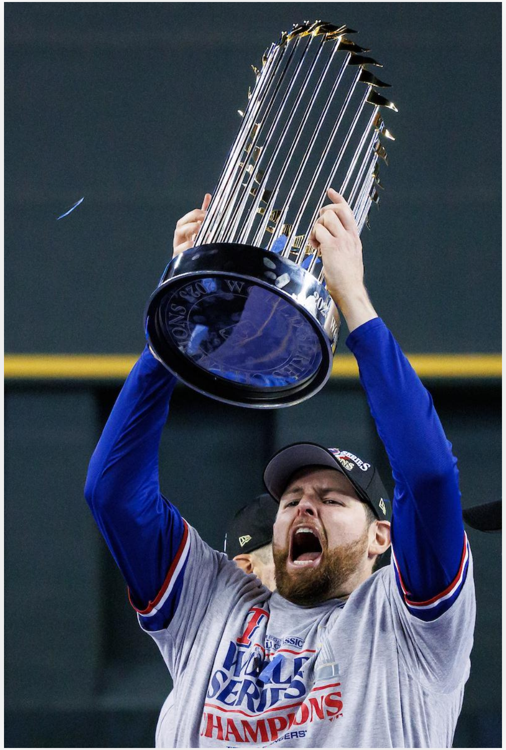
Prev Pitch	FF	SI	FC	SL	CU	CH
FB	37%	5%		16%	37%	5%
SI	4%	37%	11%	33%	4%	11%
CT		57%	14%			29%
SL	18%	36%		32%	7%	7%
CB	23%	23%	5%	32%	9%	9%
CH	9%	18%		18%	9%	45%
Other	21%	21%	11%	21%	21%	5%



World Series Victory

Putting It All Together





World Series Victory

Putting It All Together

Questions?