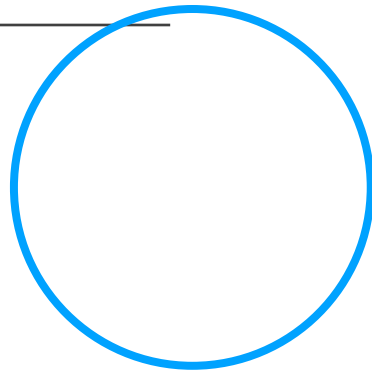


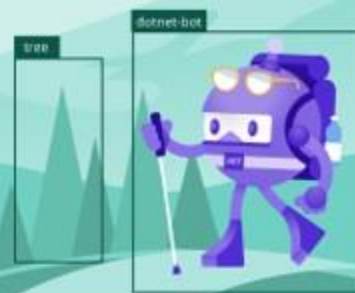
# ML Systems in Vertical Slices

---

F# & .NET INTERACTIVE 4 THE WIN



The Virtual  
**ML.NET**  
Community Conference



## Thank you to our sponsor!

Since 2002, Excella has helped clients build cutting-edge technology and transform their delivery practices

Build	Transform
Machine Learning	Modern Quality Engineering
Solution Architecture	DevOps and DataOps
Cloud Migration	Agile for Development
.NET Core Migration	Agile for Analytics/ML




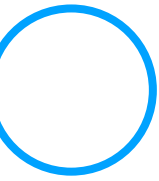
# Excella

excella.com | @excellaco



# HI Y'ALL!

- Software Engineer turned Data Scientist
  - Functional Programming Enthusiast
  - Working for HAKOM Time Series GmbH (we work with time series!)
- 



# THE INSPIRATION FOR THIS TALK

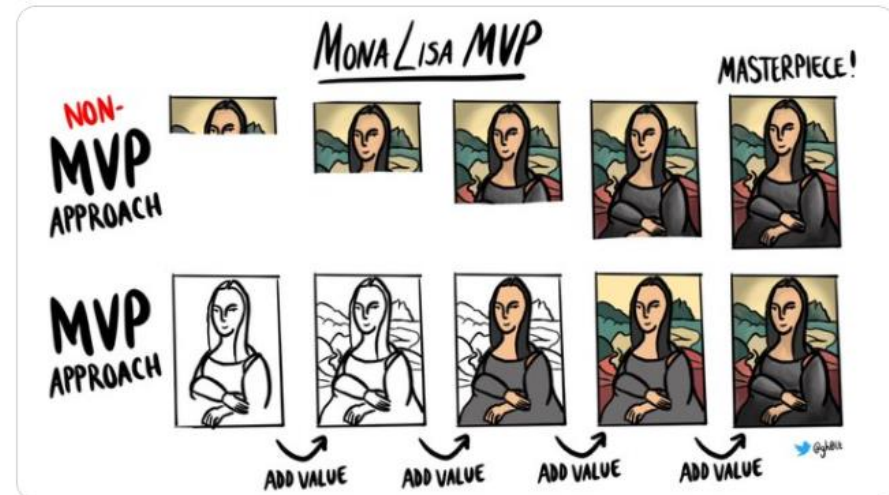
- Research and Development is a game of pitches
- ...as are Startups
- ...as is Game Development
- Launching a perfect system does seldomly happen
- Presenting a limited prototype gets feedback and wins the grant money



Greg Holt  
@gh0lt



#DeliverCon2021 day 1 was excellent! A great point raised on conveying MVP and how the skateboard/car analogy doesn't quite cut it - someone suggested the Mona Lisa instead. It really stuck with me so I took a stab at drawing it.



12:02 PM · Apr 29, 2021



136 8 Share this Tweet

# WE THINK SYSTEMS, NOT MODELS



- Start with a real problem
- Develop a hypothesis
- Collect enough data for prototype
- Sample and transform data
- Build and evaluate model
- Deploy model
- Test model with consumer and evaluate
- Rinse & Repeat

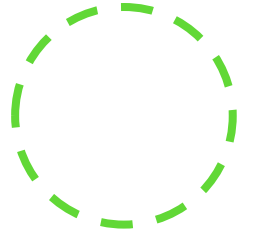
Source: <https://xkcd.com/1838/>



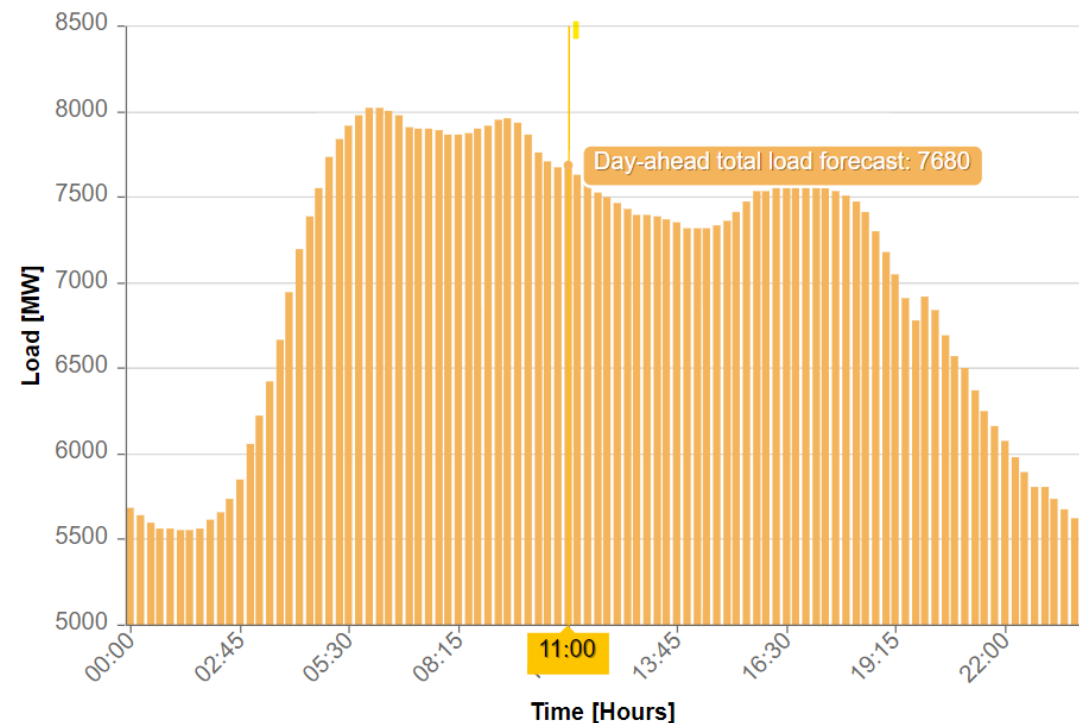
# The Problem Statement

PLEASE START HERE

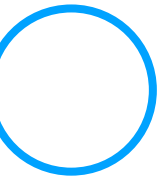
# WE WANT TO FORECAST ENERGY LOAD



- As an Energy Provider...
- I want to forecast energy load...
- At least a day ahead at an hourly rate
- In order to buy/produce the right amount at the best price



Source: <https://transparency.entsoe.eu/load-domain/r2/totalLoadR2/show>



A decorative graphic featuring two thick pink arcs that curve from the top corners towards the bottom. A dashed orange circle is in the top left, a dashed green circle is in the top right, and a solid blue circle is in the bottom right. Small solid circles in yellow, cyan, and magenta are also present.

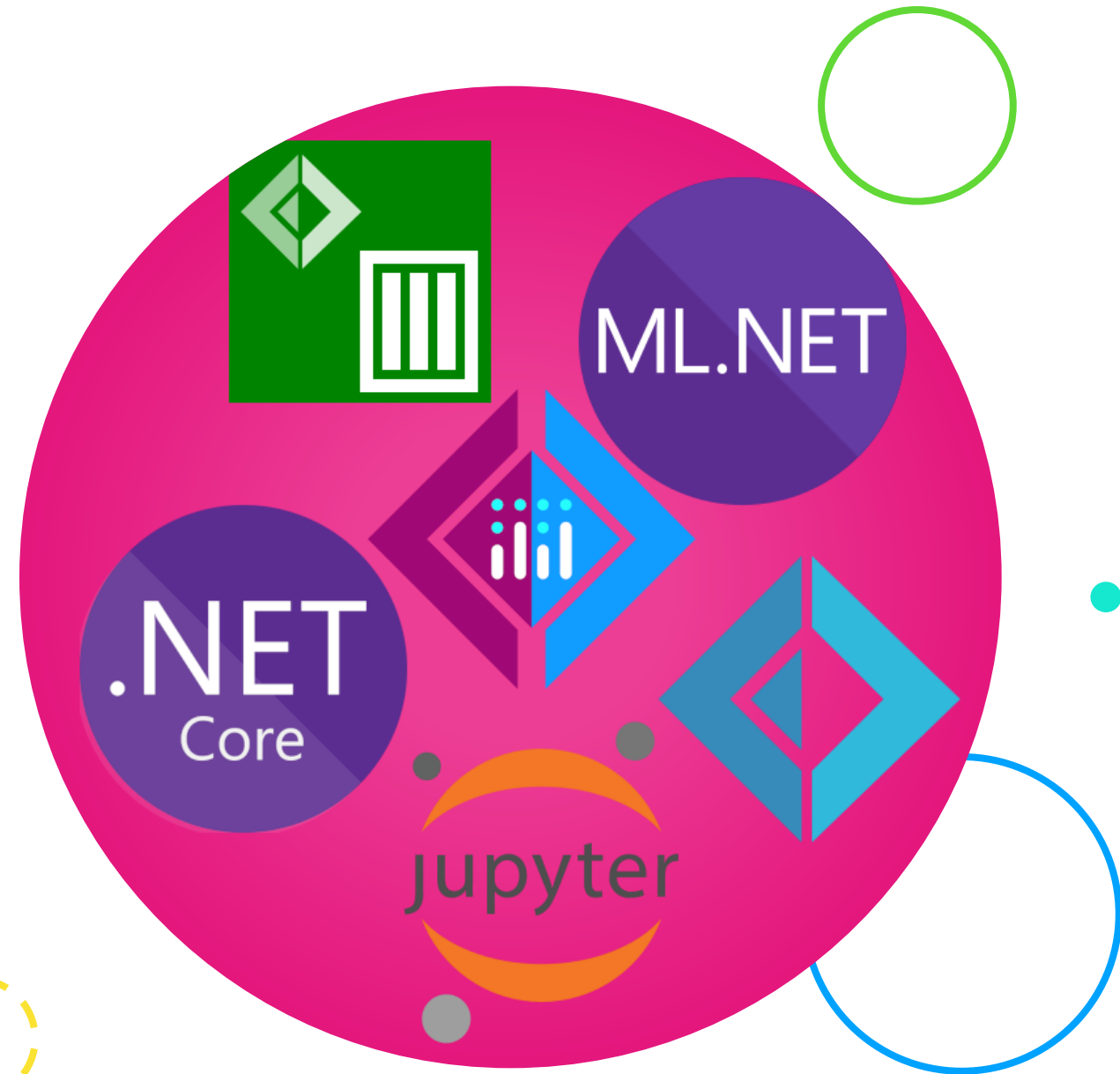
# The Tools

STANDING ON THE SHOULDERS OF  
GIANTS



# INTERACTIVE AND CONCISE

- Interactive Programming makes for fast iterations
- F# is expressive and concise
- Plotly.NET is pretty and interactive
- Deedle provides data frames and time series functions
- ML.NET for modelling
- .NET Core is a rock solid, enterprise ready™ and performs very well

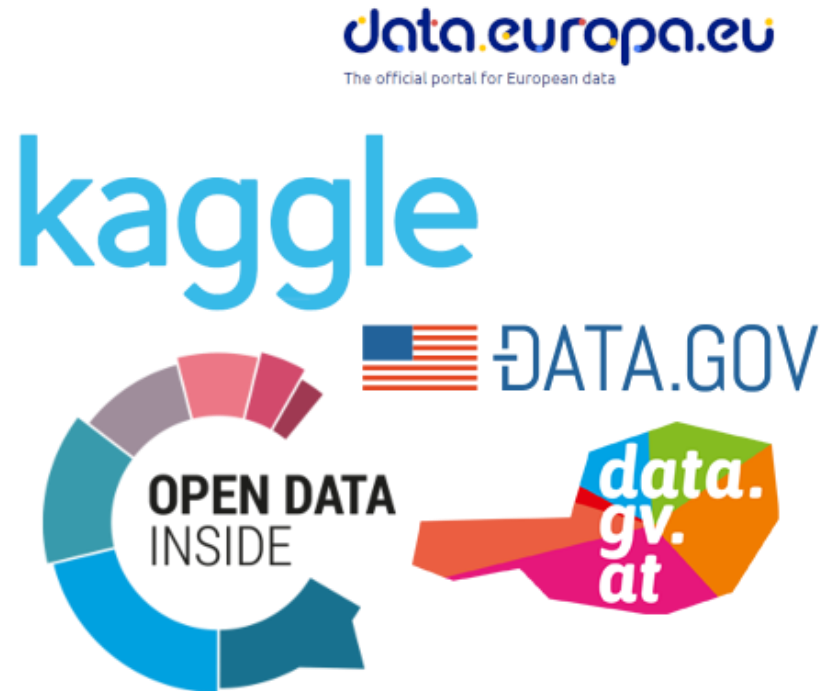




# Getting Data

BECAUSE WE NEED SOMETHING TO  
WRANGLE

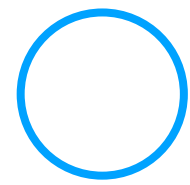
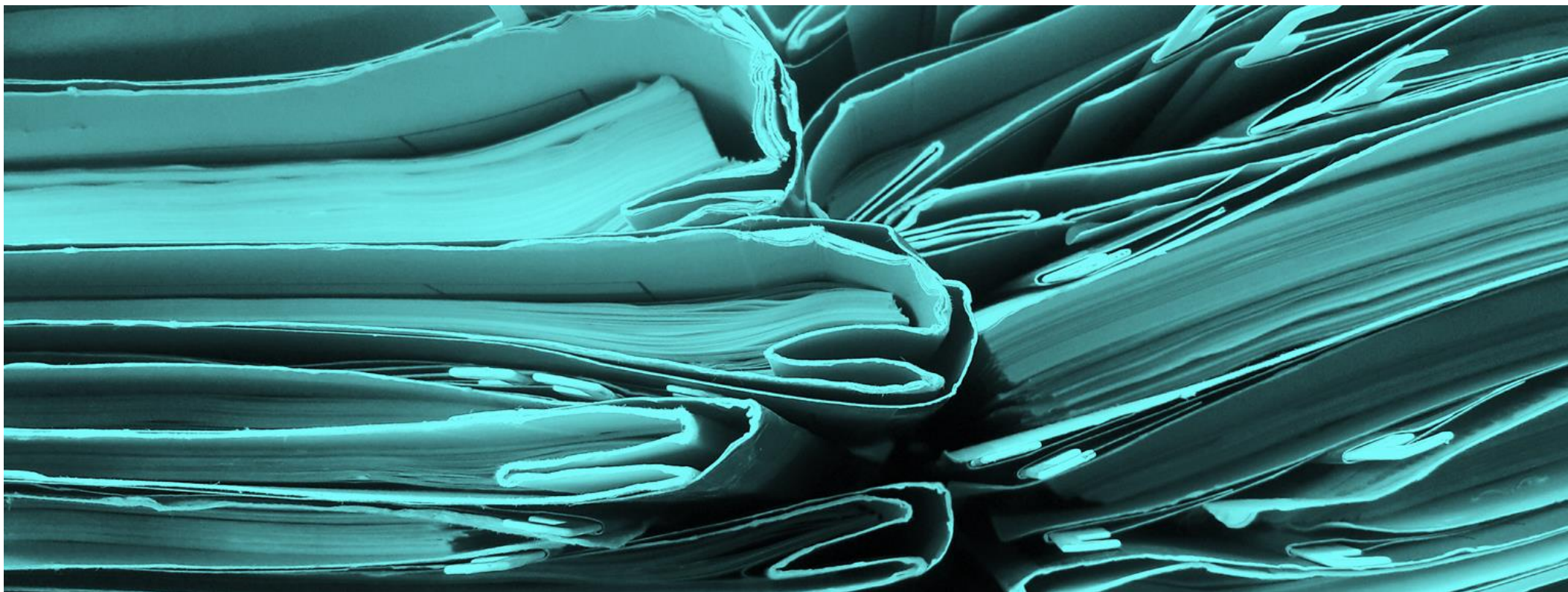
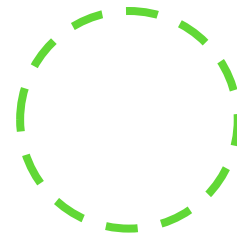
# PITCH WITH REAL DATA

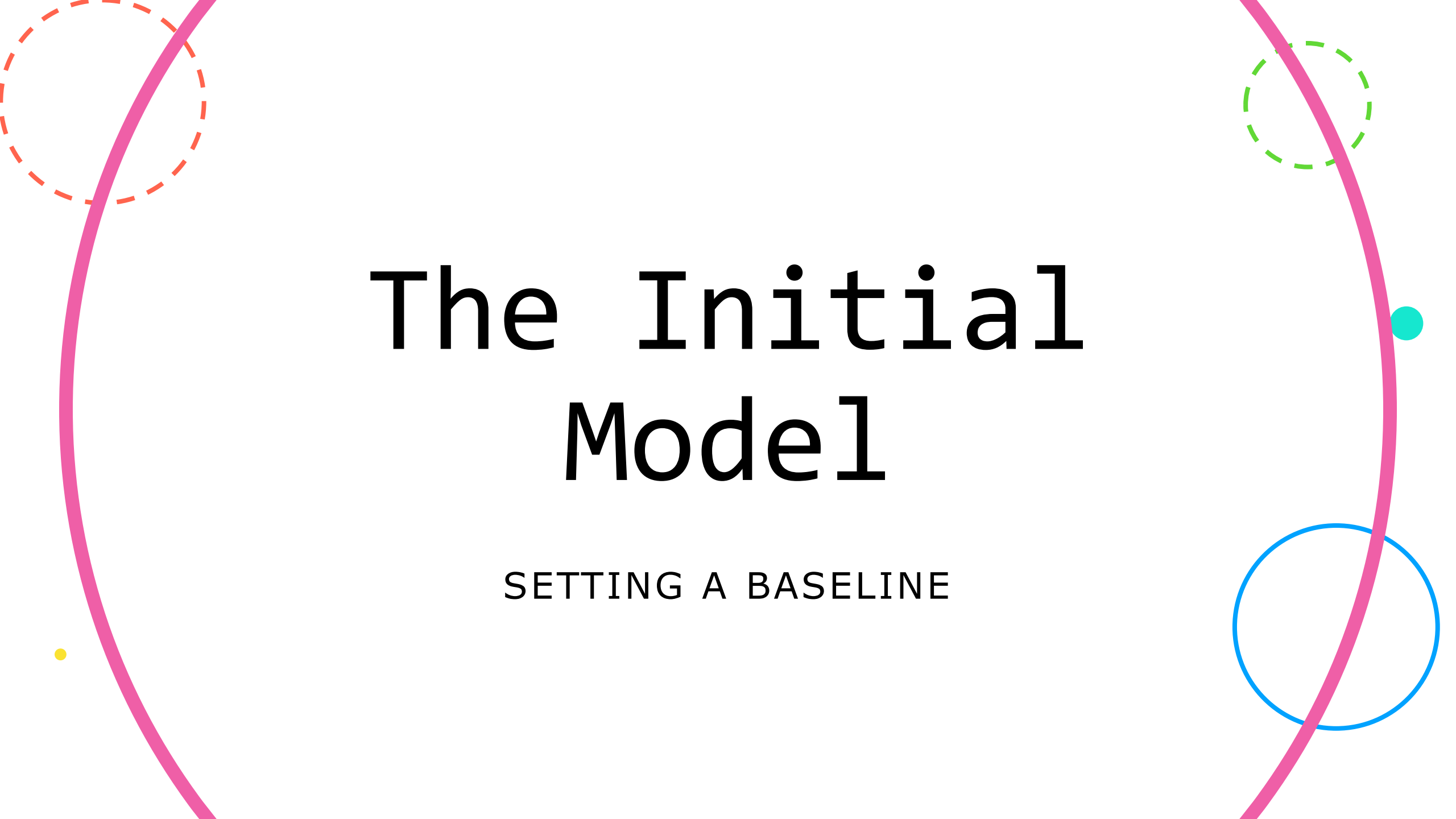


- Benchmark datasets are fun but not always useful
- There is an abundance of open data
- This data is usually not super pretty
- That is a good thing – actually
- Kaggle can be a decent first step



DEMO TIME!



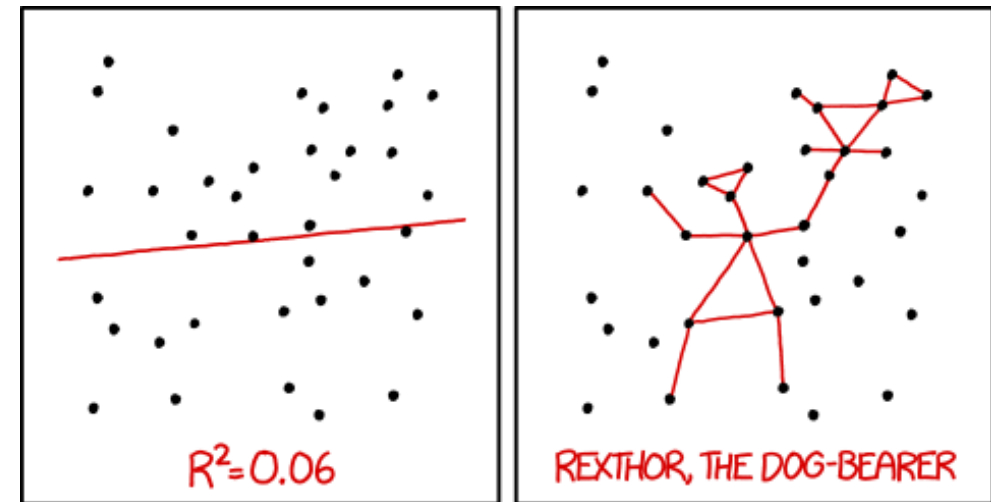


# The Initial Model

SETTING A BASELINE

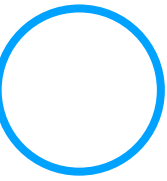
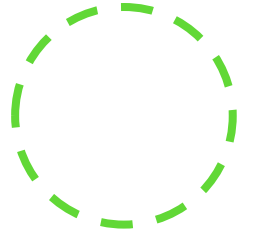
# LINEAR MODELS AND HOW TO WIELD THEM

- Linear models usually offer a good baseline
- ...either they are flexible enough to be good
- ...or they are really bad and offer a good lower bound
- For systematically stable time series linear models are a good “first throw”



I DON'T TRUST LINEAR REGRESSIONS WHEN IT'S HARDER TO GUESS THE DIRECTION OF THE CORRELATION FROM THE SCATTER PLOT THAN TO FIND NEW CONSTELLATIONS ON IT.

DIRECTLY INTO THE DEMO WE GO!





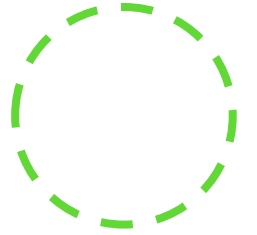
# Deploy Or It Did Not Happen

THE THING THAT IS REALLY EASY IN  
.NET

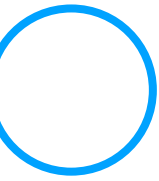




# YOU FIT A MODEL, AND NOBODY CALLS IT...

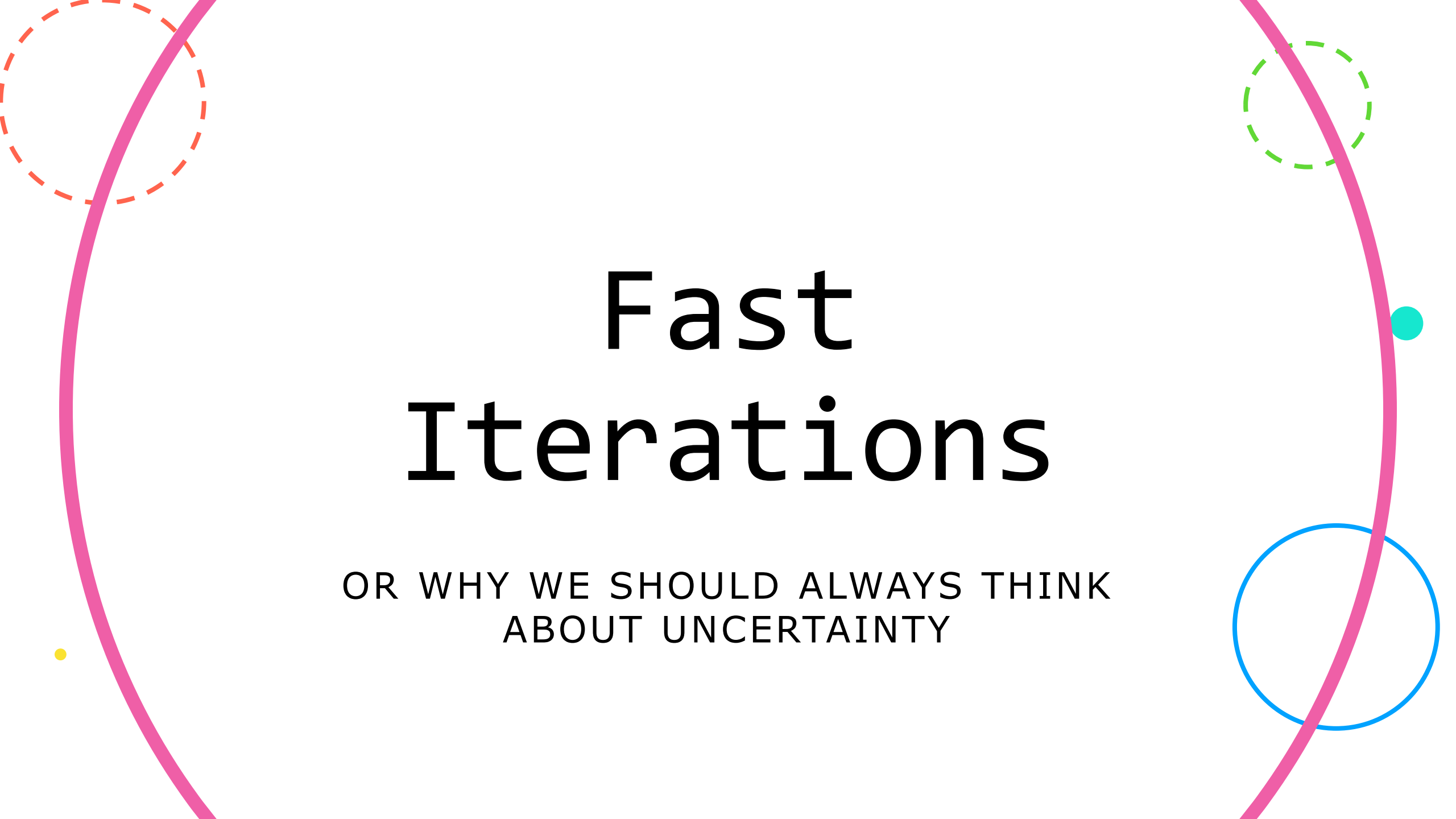


- ...does it really exist?
- .NET isn't primarily known for Data Science or ML
- But we are brilliant at getting stuff into production
- An MVP needs to be usable if you want to pitch it
- ...even if the usage is very limited



YOU GUESSED IT... HERE WE GO!



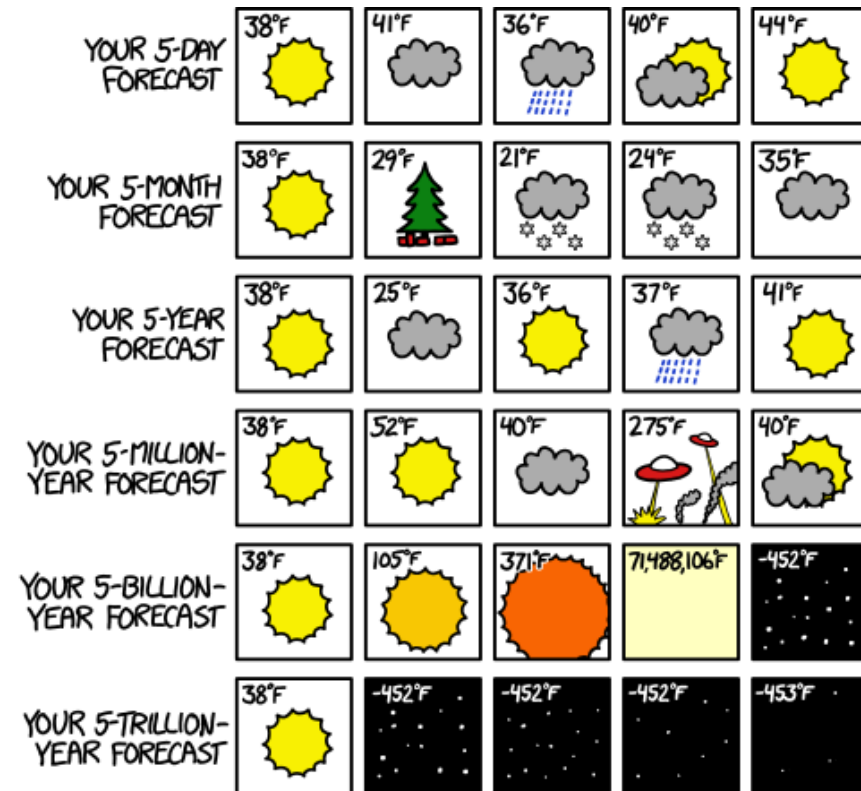


# Fast Iterations

OR WHY WE SHOULD ALWAYS THINK  
ABOUT UNCERTAINTY

# ALL FORECASTS ARE WRONG...

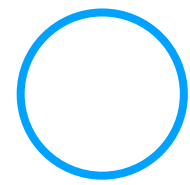
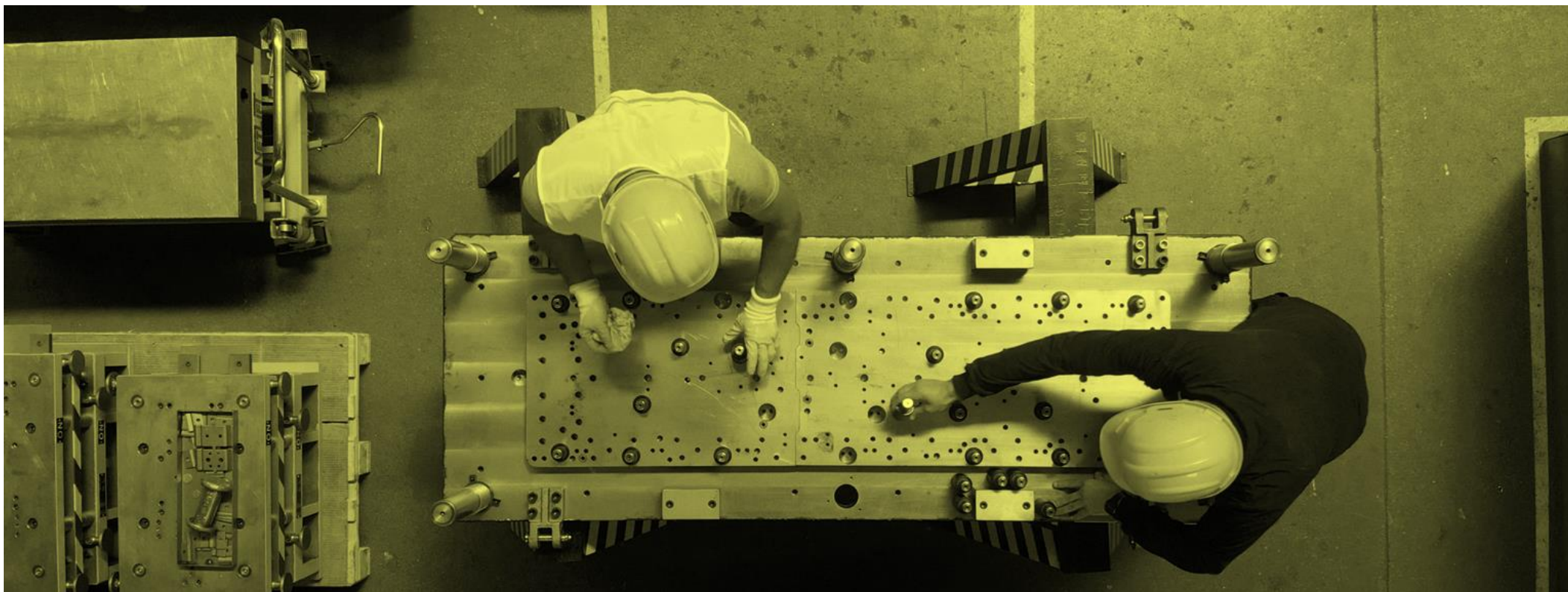
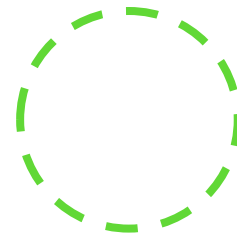
- ...but some are useful
- Currently we are far off the mark
- Worse than that: we don't factor in uncertainty
- ML.NET offers more advanced forecasting models
- As we only have one predictor, we can use Singular Spectrum Analysis



Source: <https://xkcd.com/1606/>



# BACK TO WORK!







# Where To Go From Here

YOUR SYSTEM'S LIFE AFTER THE PITCH



# POSSIBLE NEXT STEPS

- Implement experiment tracking
  - MLOps.NET
- Get more domain knowledge
- Try new models
  - Hierarchical Time Series Forecasting
  - Recurrent Neural Networks
- Get more (and better) data



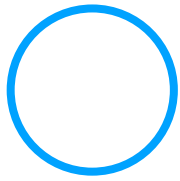


# SHAMELESS PLUG!



# HAKOM


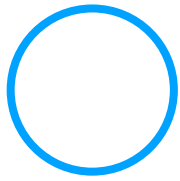
— TIME SERIES —

- SQLite and CSV doesn't scale
  - Manage hundreds of thousands of Time Series
  - Don't worry about common Time Series transformation tasks
  - Extras like auditing, meta data, compression, etc.
  - Years of experience in the field
  - Someone you can call if you have a problem
- 





# USEFUL COMMUNITY LINKS

- [HAKOM YouTube Channel](https://www.youtube.com/channel/UCMxvNphrg6qdYV280i1awxw)  
<https://www.youtube.com/channel/UCMxvNphrg6qdYV280i1awxw>
  - [The .NET Discord](https://aka.ms/dotnet-discord)  
<https://aka.ms/dotnet-discord>
  - [The FSharp Foundation Slack](https://fsharp.org/guides/slack/)  
<https://fsharp.org/guides/slack/>
  - [The FsLab GitHub Discussions](https://github.com/fslaborg/FsLab/discussions)  
<https://github.com/fslaborg/FsLab/discussions>
- 
- 



# GET IN TOUCH!



**LinkedIn**

<https://www.linkedin.com/in/gregor-beyerle/>



**Twitter**

@GBeyerle



**Homepage**

<https://www.gregorbeyerle.me/>



**GitHub**

<https://github.com/WalternativE>

