An Optimal Portfolio of Pokemon

Data Science > Business Intelligence

Vincent D. Warmerdam

GoDataDriven + koaning.io + @fishnets88

Who is this guy

- Vincent D. Warmerdam
- data guy @ GoDataDriven
- from **Amsterdam**
- avid Python, R and js user.
- give open sessions in R/Python
- minor user of scala, julia, clang.
- blog: koaning.io
- twitter: @fishnets88
- hobby: data hacking (video) games

Pokemon: A Game

python, javascript, R, Rcpp, dplyr, ggplot, ditto, scala, shiny, mongo, html5, bootstrap, git, d3, leaflet, sawk, pandas, feebas, numpy, scikit, nltk, crebase, juypter, h2o, onyx, lodash, tensorflow, docker, django, flask, neo4j, vulpix, selenium, node, hadoop, hoopa, impala, spark, azurill, ansible, hadoop, mapreduce

Pokemon: A Game

python, javascript, R, Rcpp, dplyr, ggplot, ditto, scala, shiny, mongo, html5, bootstrap, git, d3, leaflet, sawk, pandas, feebas, numpy, scikit, nltk, crebase, juypter, h2o, onyx, lodash, tensorflow, docker, django, flask, neo4j, vulpix, selenium, node, hadoop, hoopa, impala, spark, azurill, ansible, hadoop, mapreduce

Recruiters: we expect you to understand which are pokemon and which aren't.

Pokemon: Red/Blue

Pokemon Science

Classic Nintendo Games are (Computationally) Hard

Greg Aloupis* Erik D. Demaine† Alan Guo†‡ Giovanni Viglietta§

February 10, 2015

Abstract

We prove NP-hardness results for five of Nintendo's largest video game franchises: Mario, Donkey Kong, Legend of Zelda, Metroid, and Pokémon. Our results apply to generalized versions of Super Mario Bros. 1–3, The Lost Levels, and Super Mario World; Donkey Kong Country 1–3; all Legend of Zelda games; all Metroid games; and all Pokémon role-playing games. In addition, we prove PSPACE-completeness of the Donkey Kong Country games and several Legend of Zelda games.

1 Introduction

A series of recent papers have analyzed the computational complexity of playing many different video games [1, 4, 5, 6], but the most well-known classic Nintendo games have yet to be included among these results. In this paper, we analyze some of the best-known Nintendo games of all time: Mario, Donkey Kong, Legend of Zelda, Metroid, and Pokémon. We prove that it is NP-hard, and in some cases PSPACE-hard, to play generalized versions of most games in these series. In particular, our NP-hardness results apply to the NES games Super Mario Bros. 1–3, Super Mario

Pokemon: Theory

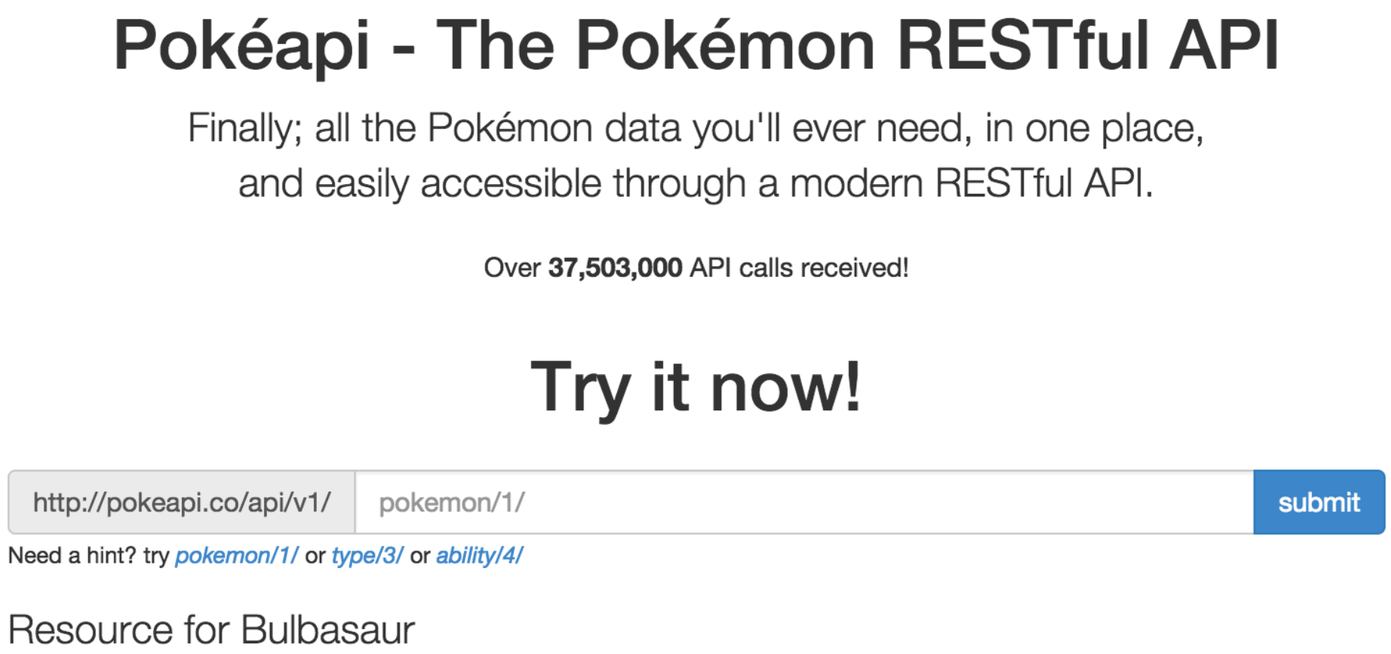
Searching fan websites gave me formulas.

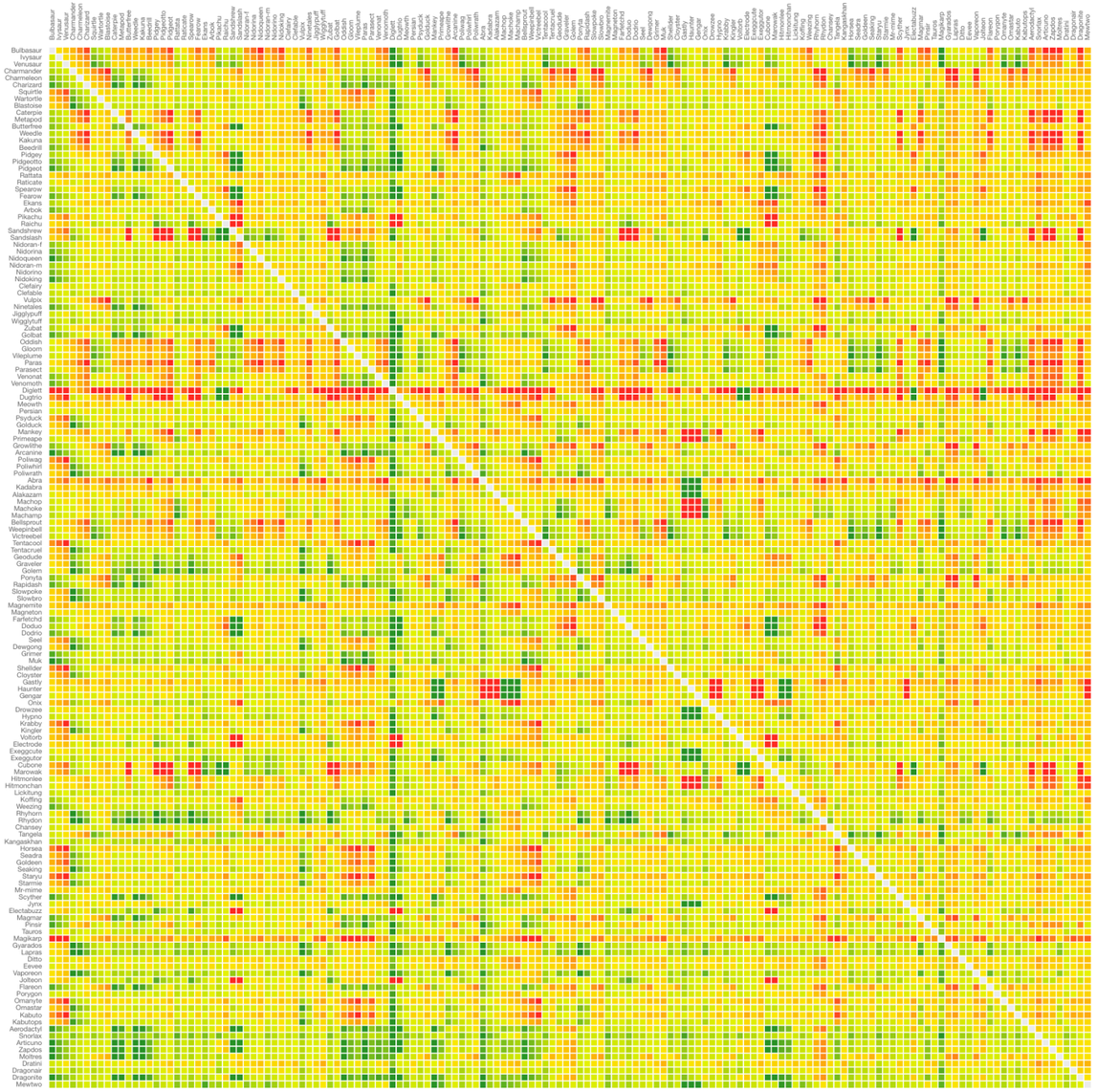
$$T_{ij} = \frac{HP_i}{DMG_{ji}}$$

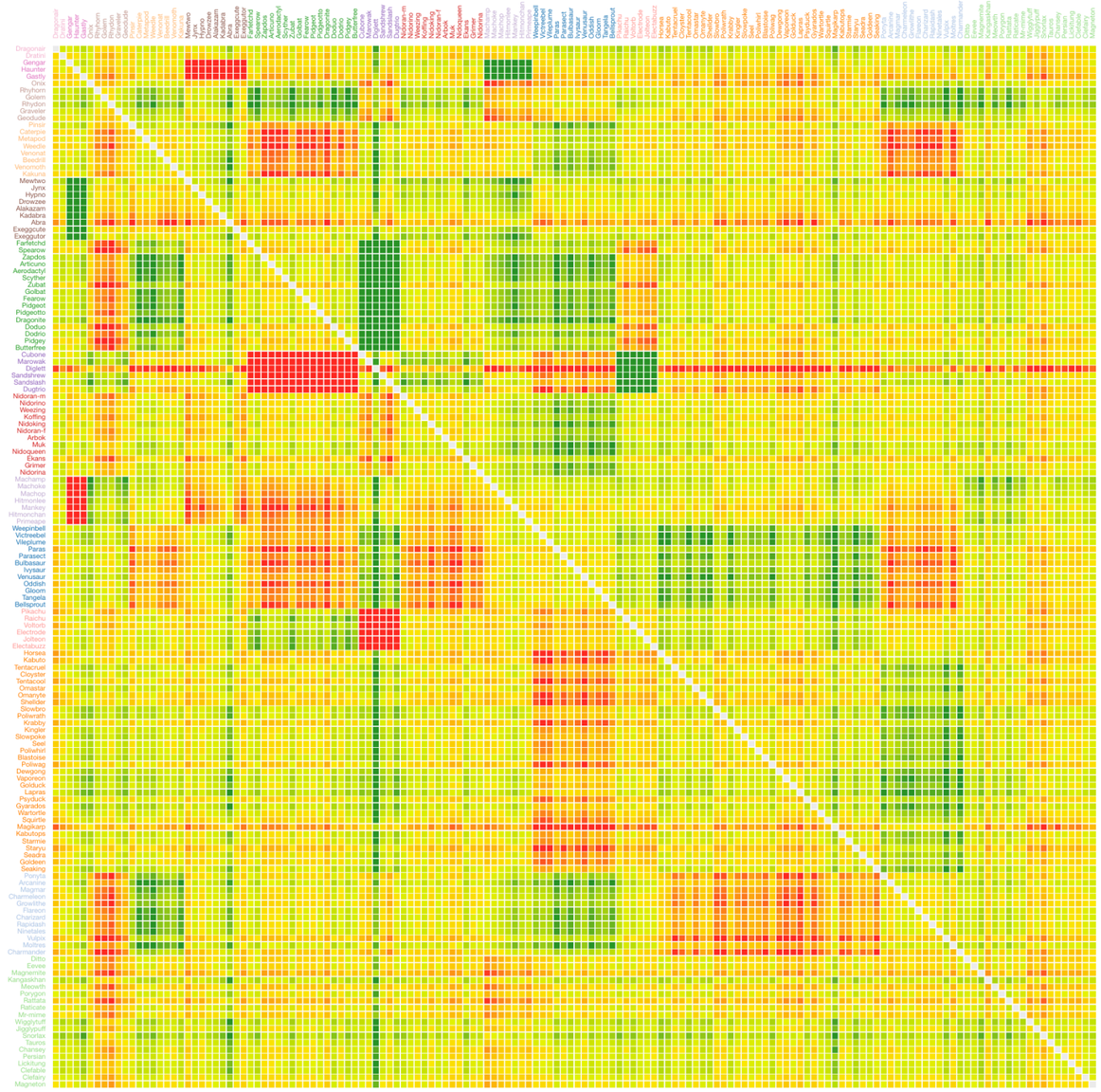
$$DMG_{ji} = \frac{2L_j + 10}{250} \times \frac{A_j}{D_i} \times w_{ji}$$

Pokemon: Data

Searching fan websites gave me data.




Bulbasaur
Ivysaur
Venusaur
Charmander
Charmeleon
Charizard
Squirtle
Wartortle
Blastoise
Caterpie
Metapod
Butterfree
Weedle
Kakuna
Beedrill
Pidgey
Pidgeotto
Pidgeot
Rattata
Raticate
Spearow
Fearow
Ekans
Arbok
Pikachu
Raichu
Sandshrew
Sandslash
Nidoran-f
Nidorina
Nidoqueen
Nidoran-m
Nidorino
Nidoking
Clefairy
Clefable
Vulpix
Ninetales
Jigglypuff
Wigglytuff
Zubat
Golbat
Oddish
Gloom
Vileplume
Paras
Parasect
Venonat
Venomoth
Diglett
Dugtrio
Meowth
Persian
Psyduck
Golduck
Mankey
Primeape
Growlithe
Arcanine
Poliwag
Poliwhirl
Poliwrath
Abra
Kadabra
Alakazam
Machop
Machoke
Machamp
Bellsprout
Weepinbell
Victreebel
Tentacool
Tentacruel
Geodude
Graveler
Golem
Ponyta
Rapidash
Slowpoke
Slowbro
Magnemite
Magneton
Farfetchd
Doduo
Dodrio
Seel
Dewgong
Grimer
Muk
Shellder
Cloyster
Gastly
Haunter
Gengar
Onix
Drowzee
Hypno
Krabby
Kingler
Voltorb
Electrode
Exeggcute
Exeggutor
Cubone
Marowak
Hitmonlee
Hitmonchan
Lickitung
Koffing
Weezing
Rhyhorn
Rhydon
Chansey
Tangela
Kangaskhan
Horsea
Seadra
Goldeen
Seaking
Staryu
Starmie
Mr-mime
Scyther
Jynx
Electabuzz
Magmar
Pinsir
Tauros
Magikarp
Gyarados
Lapras
Ditto
Eevee
Vaporeon
Jolteon
Flareon
Porygon
Omanyte
Omastar
Kabuto
Kabutops
Aerodactyl
Snorlax
Articuno
Zapdos
Moltres
Dratini
Dragonair
Dragonite
Mewtwo

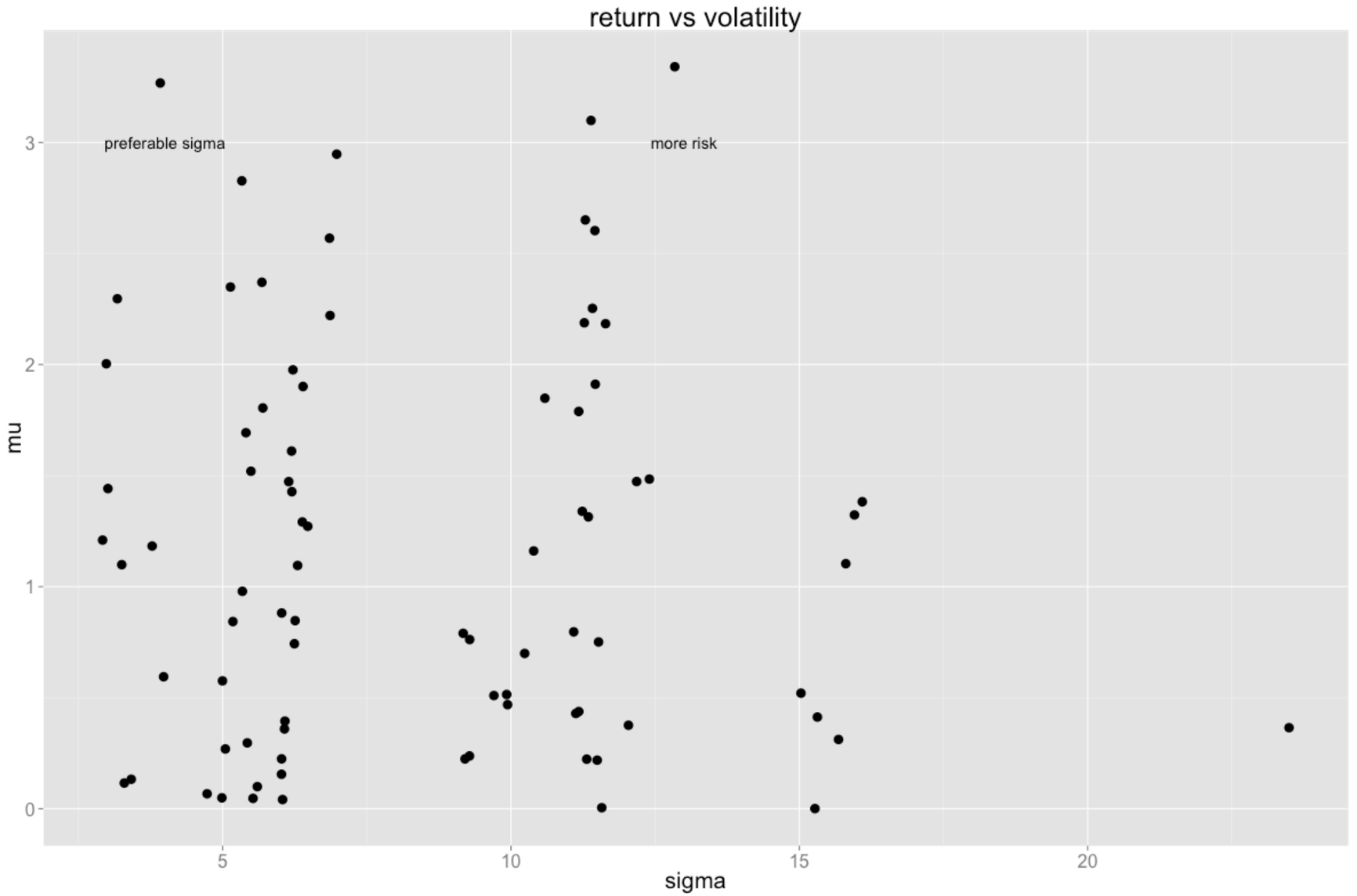

Dragonair
Dratini
Gengar
Haunter
Gastly
Onix
Rhyhorn
Golem
Rhydon
Graveler
Geodude
Pinsir
Caterpie
Metapod
Weedle
Venonat
Beedrill
Venomoth
Kakuna
Mewtwo
Jynx
Hypno
Drowzee
Alakazam
Kadabra
Abra
Exeggcute
Exeggutor
Farfetchd
Spearow
Zapdos
Articuno
Aerodactyl
Scyther
Zubat
Golbat
Fearow
Pidgeot
Pidgeotto
Dragonite
Doduo
Dodrio
Pidgey
Butterfree
Cubone
Marowak
Diglett
Sandslash
Sandshrew
Dugtrio
Nidoran-m
Nidorino
Weezing
Koffing
Nidoking
Nidoran-f
Arbok
Muk
Nidoqueen
Ekans
Grimer
Nidorina
Machamp
Machoke
Machop
Hitmonlee
Mankey
Hitmonchan
Primeape
Weepinbell
Victreebel
Vileplume
Paras
Parasect
Bulbasaur
Ivysaur
Venusaur
Oddish
Gloom
Tangela
Bellsprout
Pikachu
Raichu
Voltorb
Electrode
Jolteon
Electabuzz
Horsea
Kabuto
Tentacruel
Cloyster
Tentacool
Omastar
Omanyte
Shellder
Slowbro
Poliwrath
Krabby
Kingler
Slowpoke
Seel
Poliwhirl
Blastoise
Poliwag
Dewgong
Vaporeon
Golduck
Lapras
Psyduck
Gyarados
Wartortle
Squirtle
Magikarp
Kabutops
Starmie
Staryu
Seadra
Goldeen
Seaking
Ponyta
Arcanine
Magmar
Charmeleon
Growlithe
Flareon
Charizard
Rapidash
Ninetales
Vulpix
Moltres
Charmander
Ditto
Eevee
Magnemite
Kangaskhan
Meowth
Porygon
Rattata
Raticate
Mr-mime
Wigglytuff
Jigglypuff
Snorlax
Tauros
Chansey
Persian
Lickitung
Clefable
Clefairy
Magneton

A bit of financial maths

Here μ_i is the expected 'win' and σ is the risk of losing.

$$\mu_i = \frac{\sum_{j=1}^{N} T_{ij}}{N}$$

$$\sigma_i = \sqrt{\frac{\sum_{j=1}^{N} (T_{ij} - \mu_i)^2}{N-1}}$$

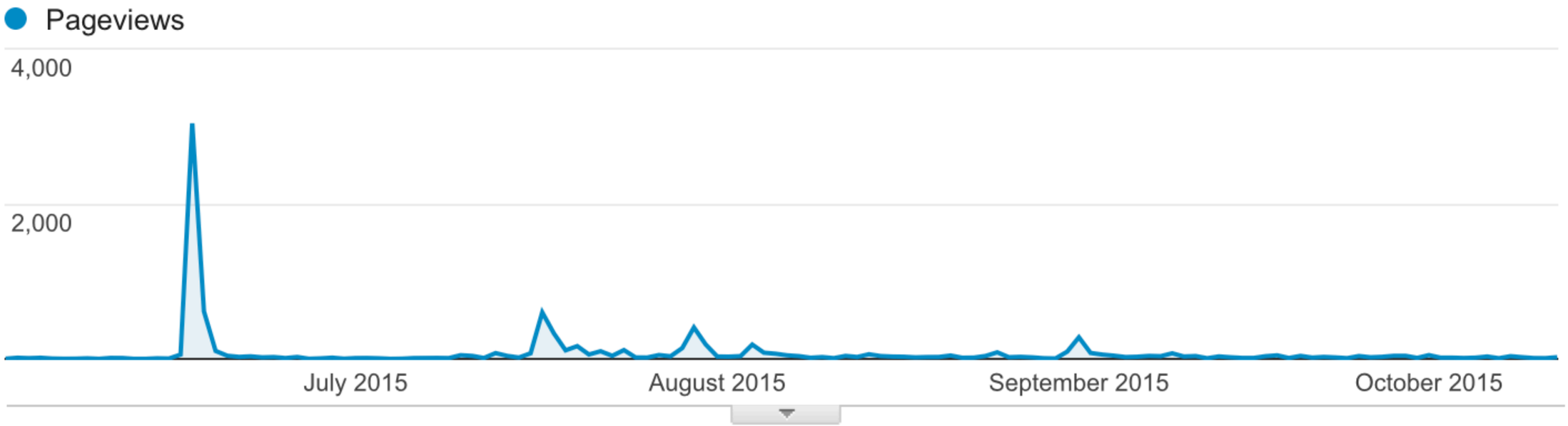

return vs volatility
preferable sigma
more risk
mu
sigma
0
1
2
3
5
10
15
20

Good portfolio of pokemon

	mu	sigma
Wigglytuff	2.296484	3.173598
Snorlax	3.267974	3.917765
Exeggutor	2.349300	5.135687
Mewtwo	2.827446	5.332813
Muk	2.370720	5.680507

And then this happend ...

Guess when the post was written



/r/pokemon/

reddit

/r/pokemon/

Feedback:

- pokemon fans did **not** agree with my modelling methods

/r/pokemon/

Feedback:

- pokemon fans did **not** agree with my modelling methods
- pokemon fans did agree that my models output made sense

/r/pokemon/

Feedback:

- pokemon fans did **not** agree with my modelling methods
- pokemon fans did agree that my models output made sense

Why this matters:

- pokemon is relatively complicated
- a good tactic without any domain knowledge

Main takeaway

- We can quickly bootstrap solutions for problems with open source tools if we have data.
- If you don't have the data, consider generating it yourself with a few assumptions or scraping it.
- I use video games as an example but obviously this approach goes beyond video games. We really can make the world (or a business) a whole lot better with data.

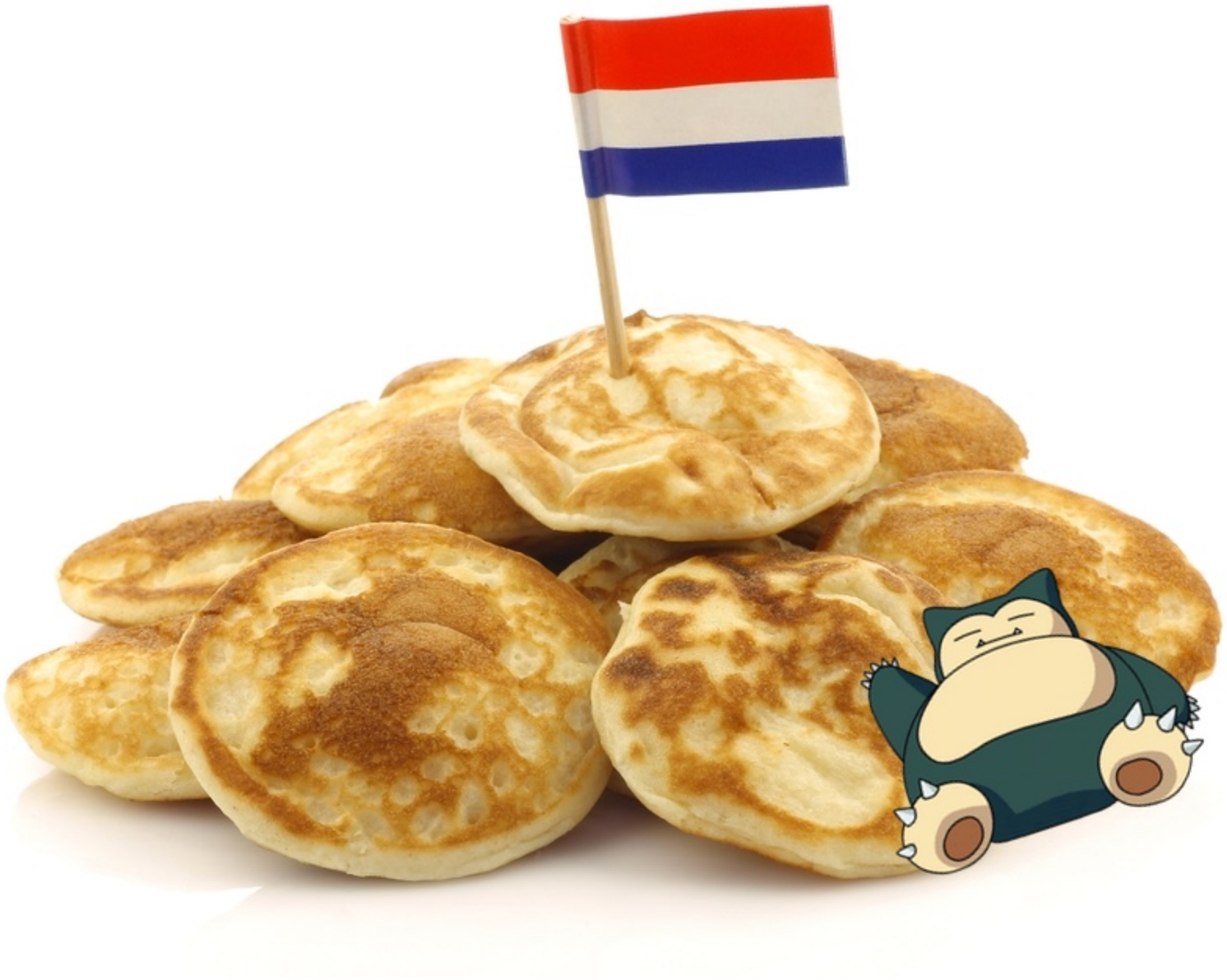
More info?

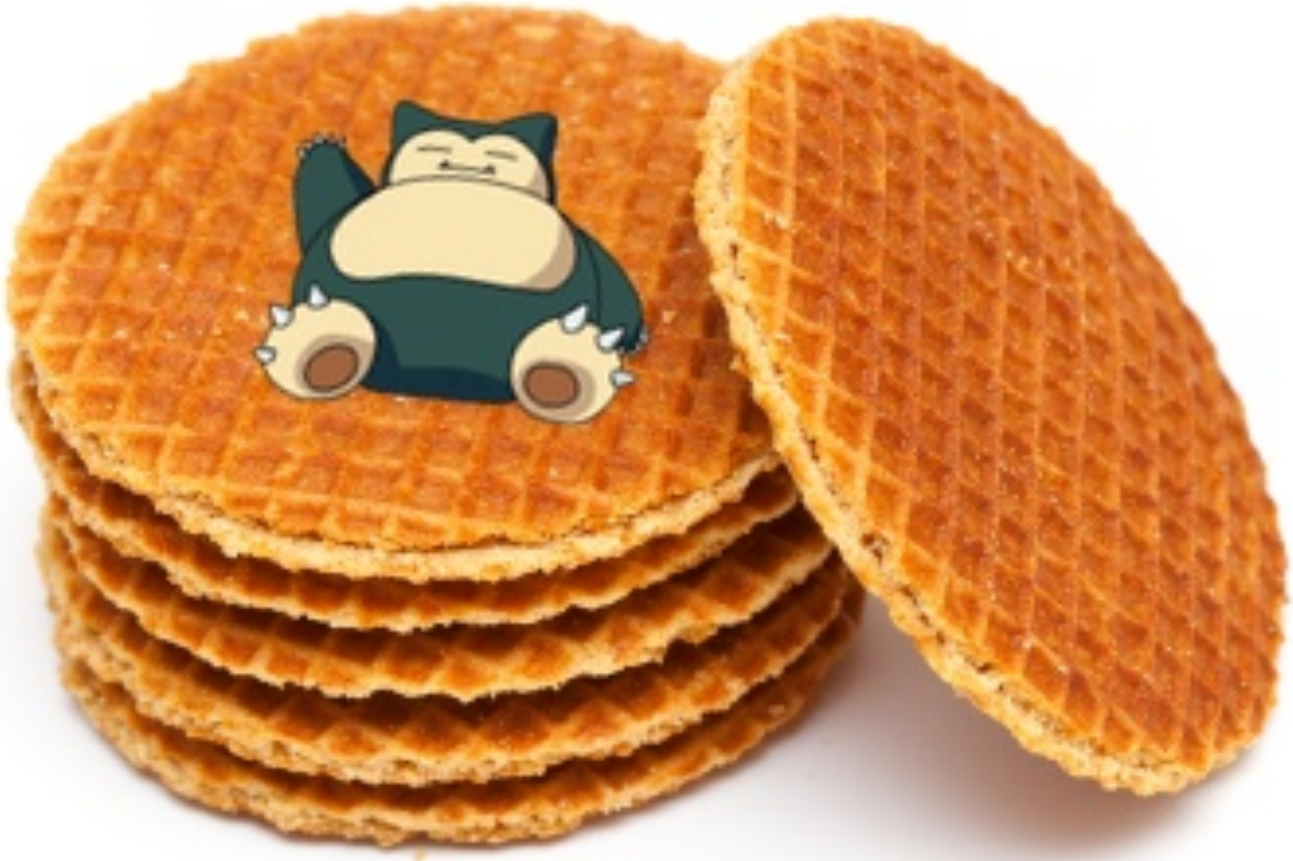
Check out the blogpost at koaning.io.

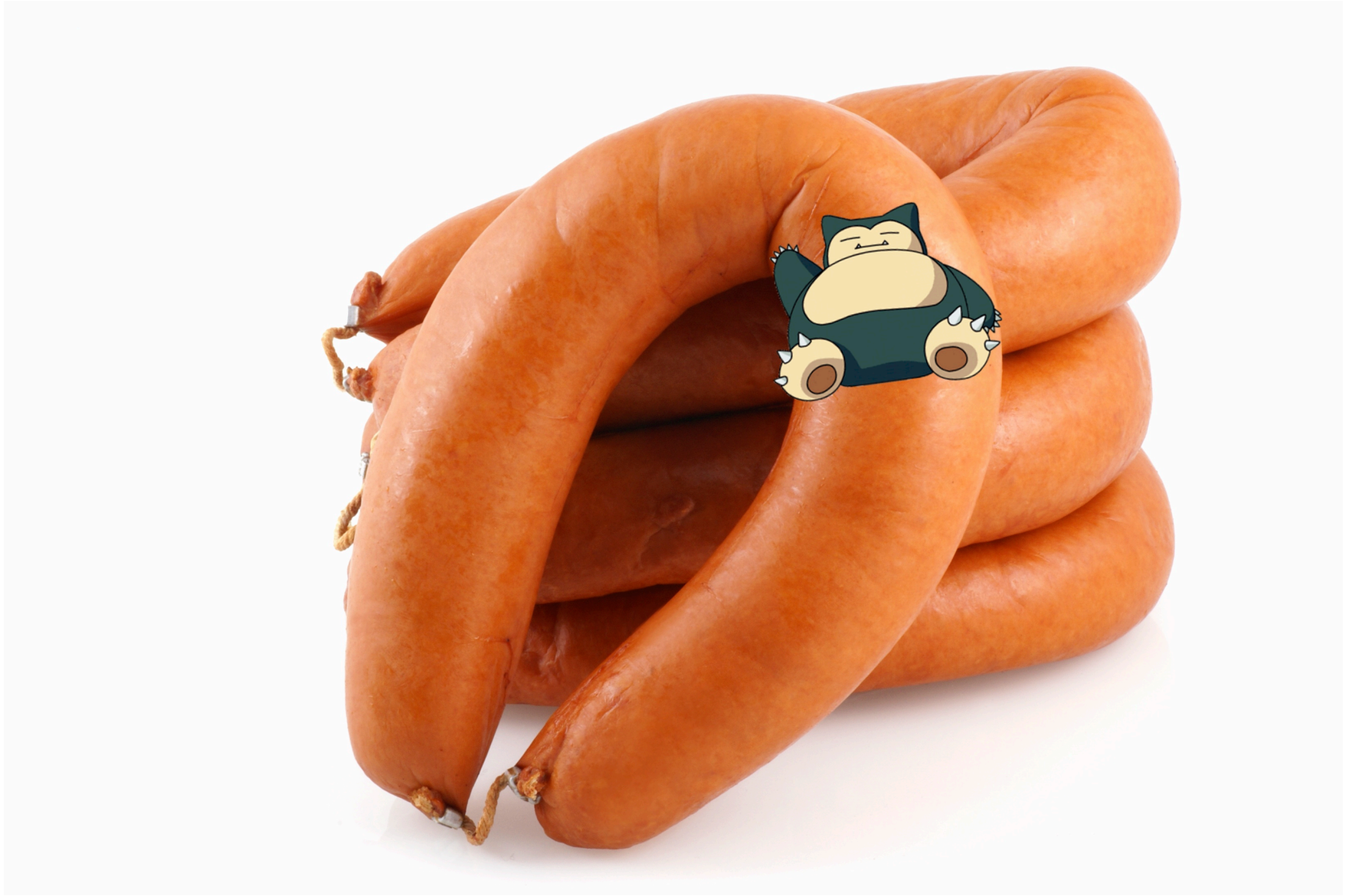
One more thing ...

... at the risk of sounding silly: que the Dutch accent.

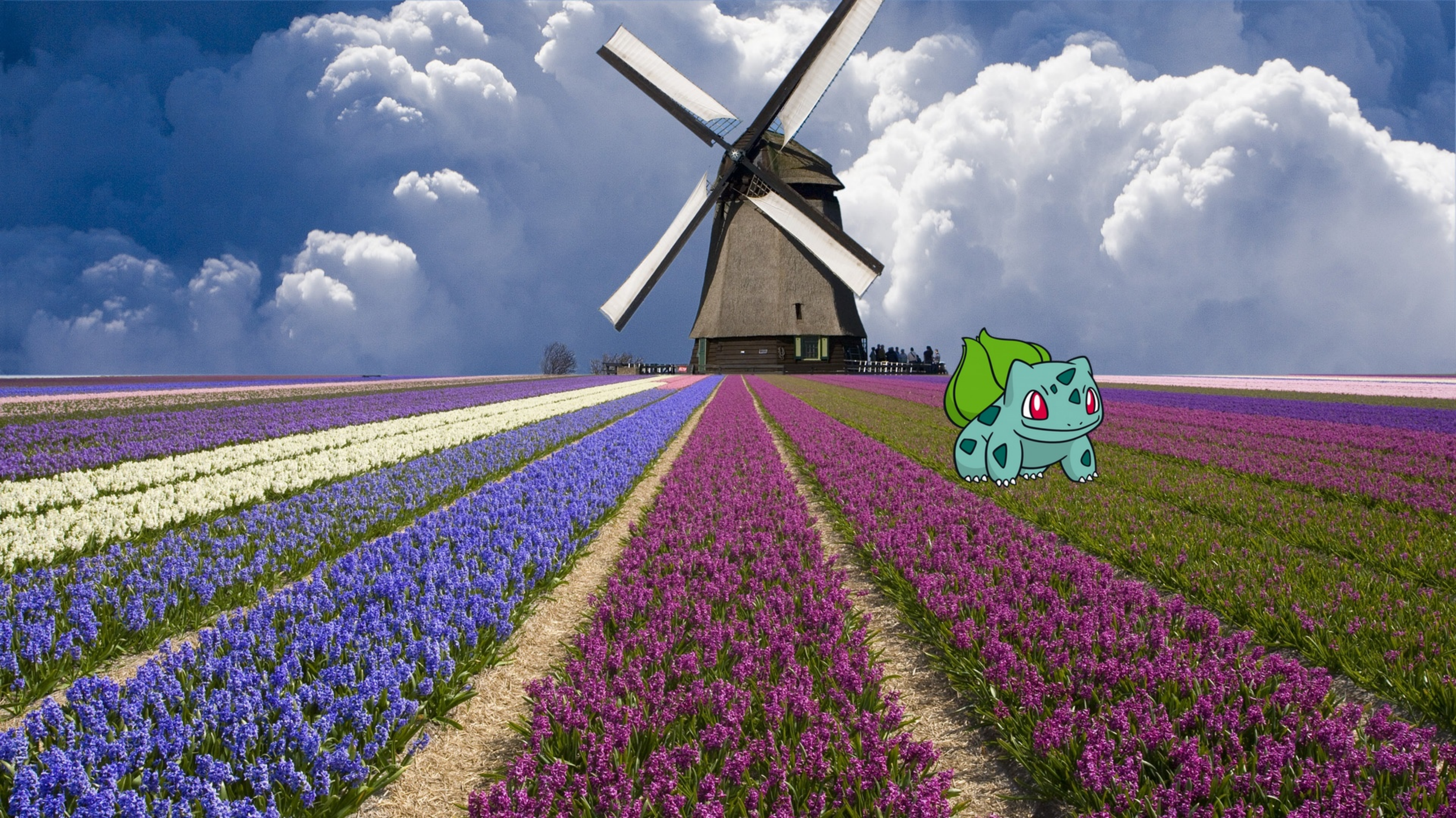

You must come to
Amsterdam!








6X
Duvel



Best Reason to Visit Amsterdam!

12-13 March 2016


Genius
TREBLE
BASS
LINE IN
VOLUME

Come to Amsterdam!

1. Did you do something awesome in 2015?
   CFP is still open (you talk, ticket = free)!
2. Do you commit to cool projects?
   We'd love to some more tutorials (also free tickets)!
3. Tickets are going fast, we can only have 300 people.
4. We'd love to see our London neighbords represented!