



▶ **Pranesh Ramjith** - BSc Hons CompSc, NDIP PM  
UTILITY ANALYTICS BUSINESS AREA MANAGER  
Energy Solutions & Analytics  
T +27 87 803 3925  
M +27 82 991 0274  
E Pranesh.Ramjith@NEXTEC.co.za

Building 13, Woodmead Estate, 1 Woodmead Drive,  
Woodmead, 2191

www.NEXTEC.co.za  
www.NEXTEC-ESA.co.za

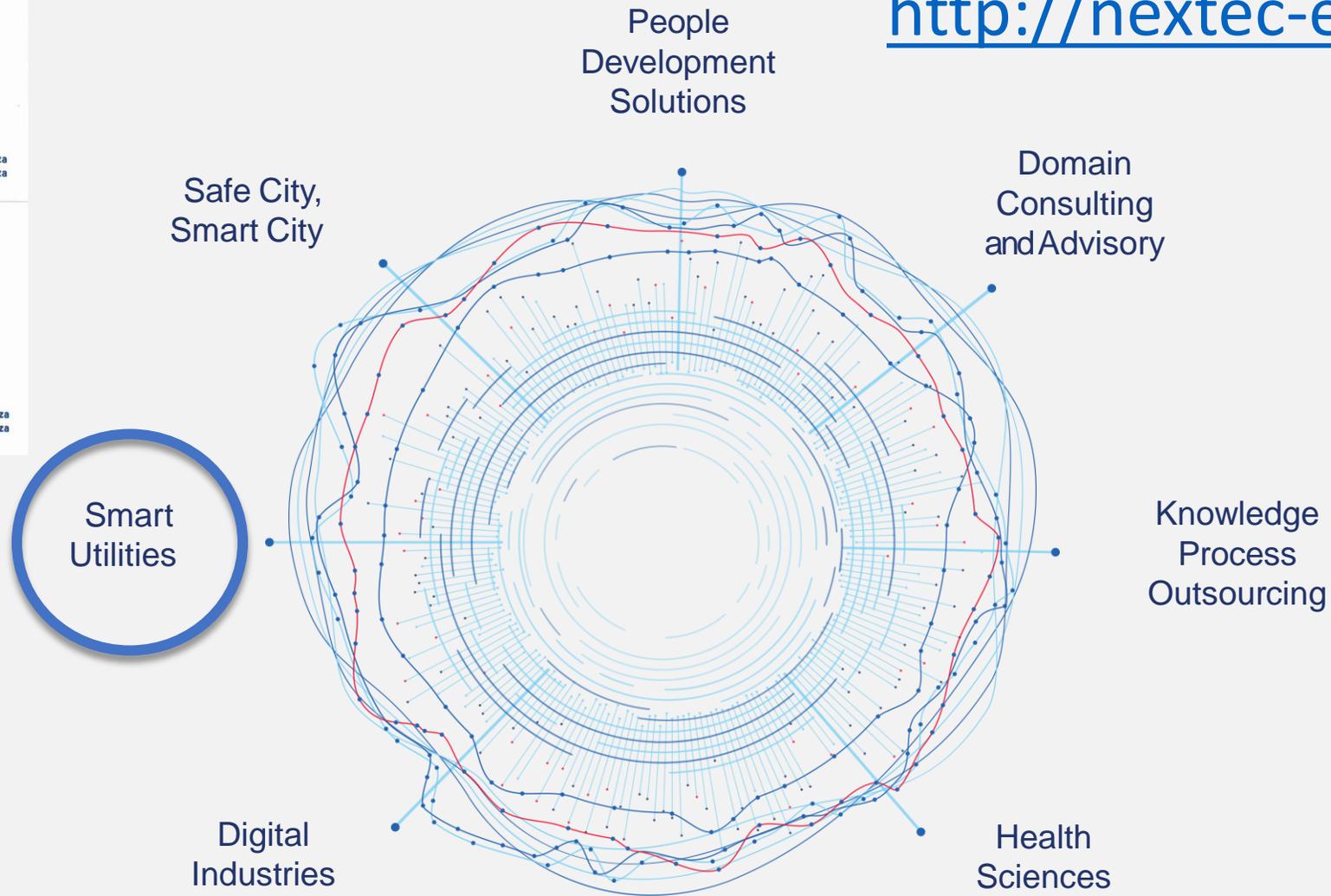


▶ **Jacques Booyesen**  
DATA SCIENTIST  
Energy Solutions & Analytics  
T +27 87 803 3927  
M +27 83 508 5567  
E Jacques.Booyesen@NEXTEC.co.za

Building 13, Woodmead Estate, 1 Woodmead Drive,  
Woodmead, 2191

www.NEXTEC.co.za  
www.NEXTEC-ESA.co.za

<http://nextec-esa.co.za>



# Data Visualisation for Load Research

Jacques Booyesen – Data Scientist NexTec -



# ► What is Data Visualisation

“Visualisation is the use of interactive visual representations of data to **amplify cognition**.”

“The purpose of [information] Visualisation is to **amplify cognitive performance**, not just to create interesting pictures. Information Visualisations should do for the mind what automobiles do for the feet.”

“[Information] Visualisation promises to help us **speed our understanding and action** in a world of increasing information volumes.”

# ► Visualisation, Exploration versus Explanation?

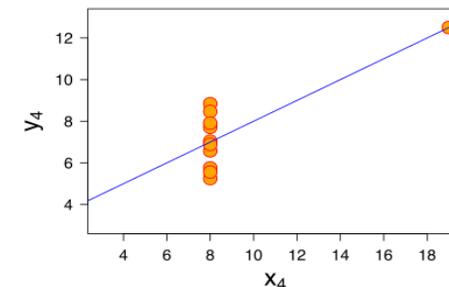
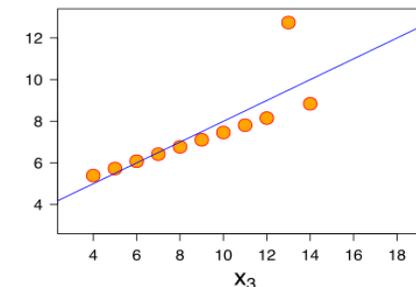
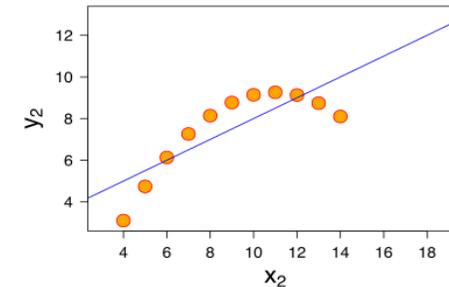
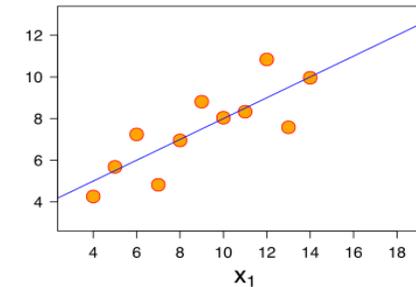
- Two categories of data Visualisation: *exploration* and *explanation*
- The two serve different purposes, mostly for pre and post modeling
- Four data sets that have nearly identical simple descriptive statistics
- But very different distributions and appear very different when Visualised.

Upper Left Plot	
x	y
10	8.04
8	6.95
13	7.58
9	8.81
11	8.33
14	9.96
6	7.24
4	4.26
12	10.84
7	4.82
5	5.68
Mean	7.50
Correl.	0.82

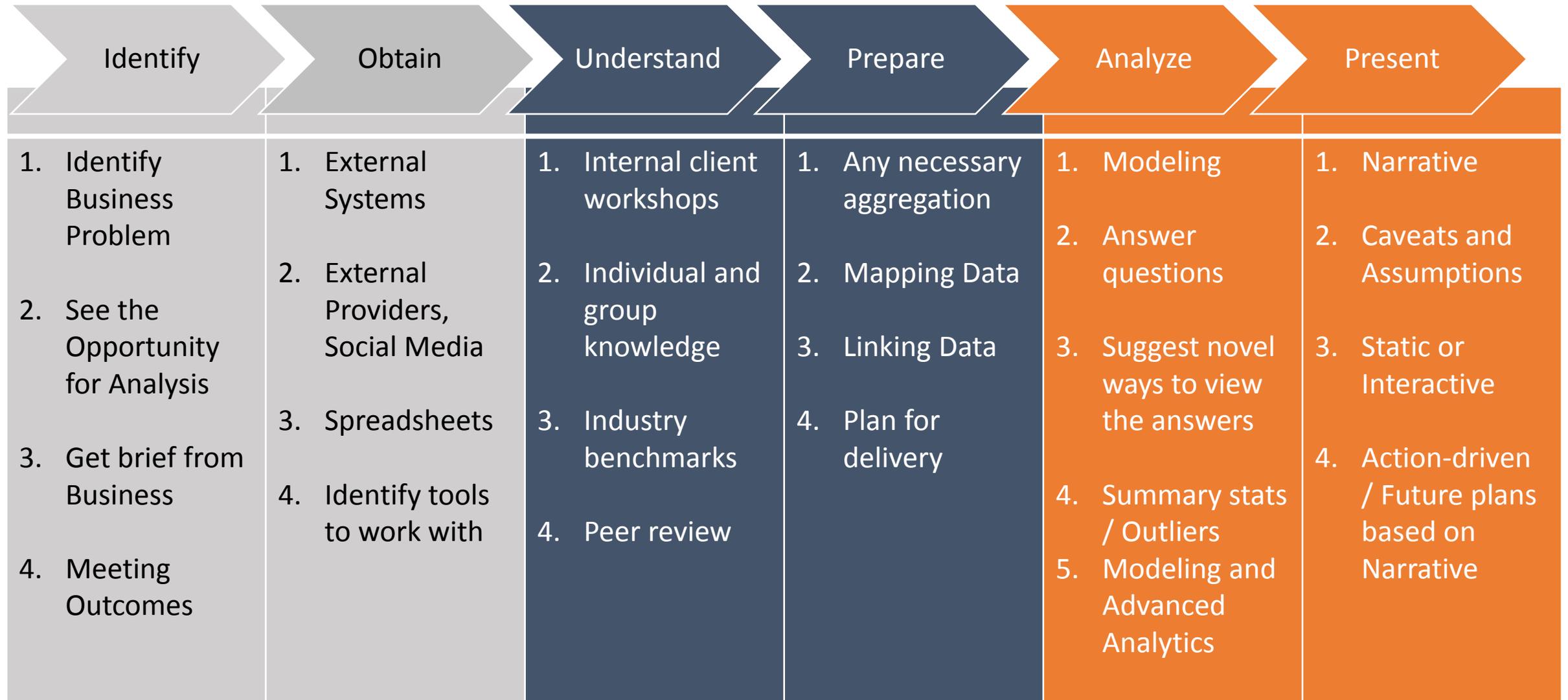
Upper Right Plot	
x	y
10	9.14
8	8.14
13	8.74
9	8.77
11	9.26
14	8.1
6	6.13
4	3.1
12	9.13
7	7.26
5	4.74
Mean	7.50
Correl.	0.82

Lower Left Plot	
x	y
10	7.46
8	6.77
13	12.74
9	7.11
11	7.81
14	8.84
6	6.08
4	5.39
12	8.15
7	6.42
5	5.73
Mean	7.50
Correl.	0.82

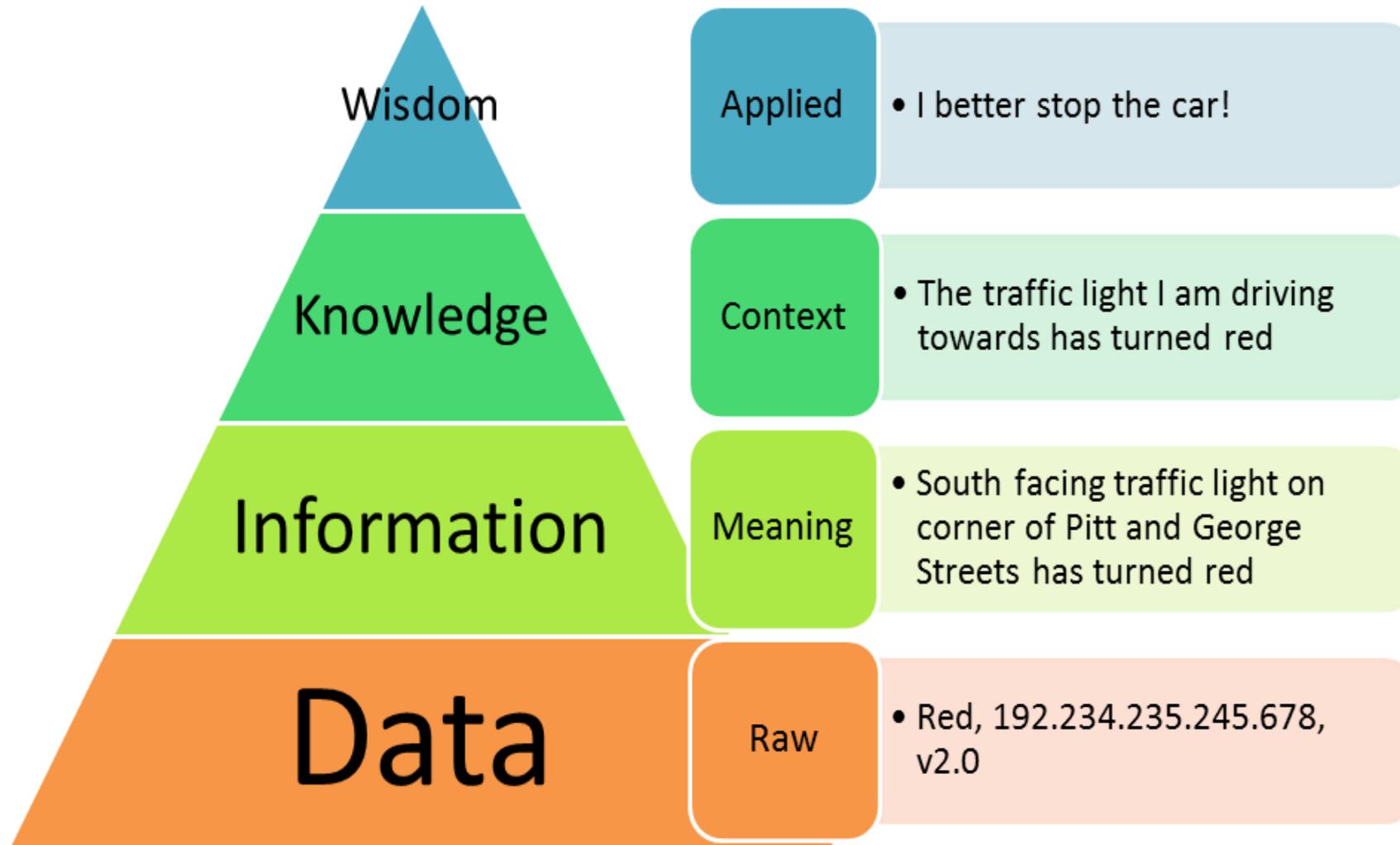
Lower Right Plot	
x	y
8	6.58
8	5.76
8	7.71
8	8.84
8	8.47
8	7.04
8	5.25
19	12.5
8	5.56
8	7.91
8	6.89
Mean	7.50
Correl.	0.82



# ▶ Process of Data Visualisation

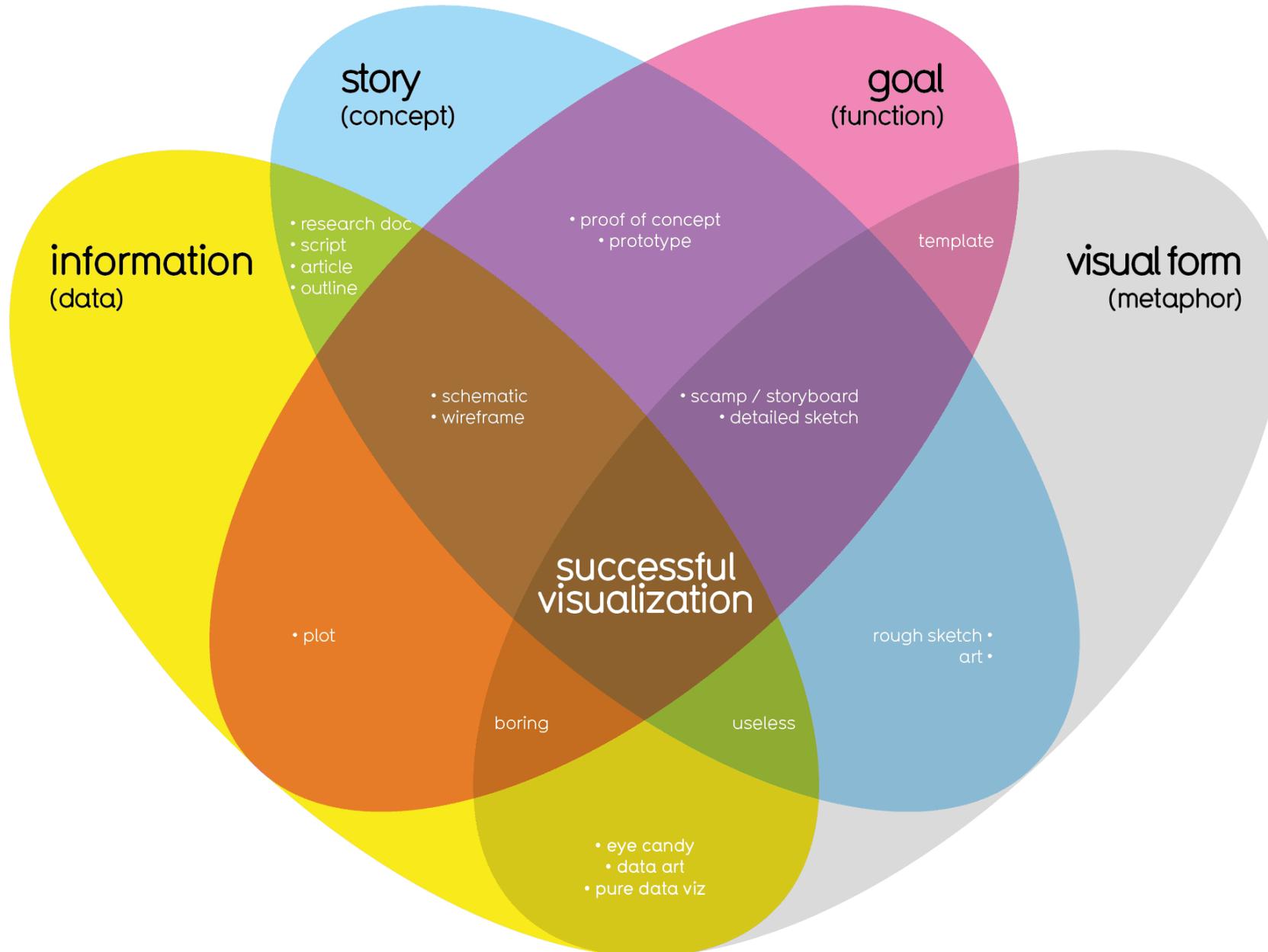


# ► The Value of Good Visualisation:



- Each layer lays a foundation for the next.
- The raw data and it's quality at the bottom is very important
- Meaning is given by linking, aggregating, mapping and cleaning the data
- By modeling the information and applying it to the context of the problem domain knowledge is gained
- This followed by **successful Visualisation**, iteration and review can assist in the journey to Wisdom and Action

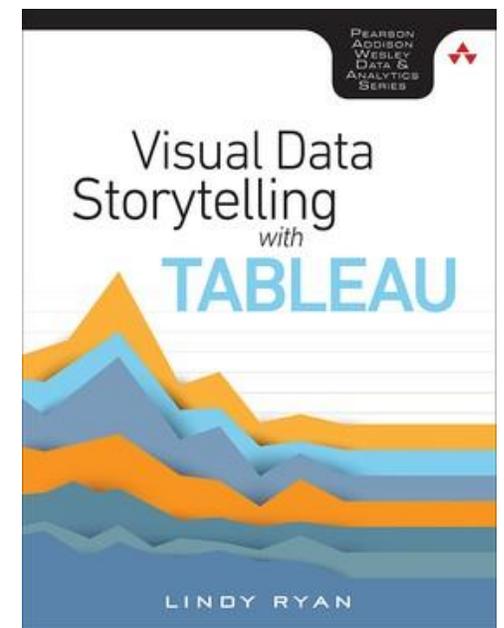
# ► What makes a Good Visualisation?



- All four elements in this graphic seem essential.
- With just two, you get something prototypical and sketchy – which is okay if you’re at an early stage.
- With three, the end result is strangely lacking.
- Combine **information & function & visual** form without story, you get “**boring**”.
- Something that looks good but isn’t that interesting.
- Combine **visuals, information & story** without considering functionality and your goal
- you get something **useless**.

# ▶ Story (concept)

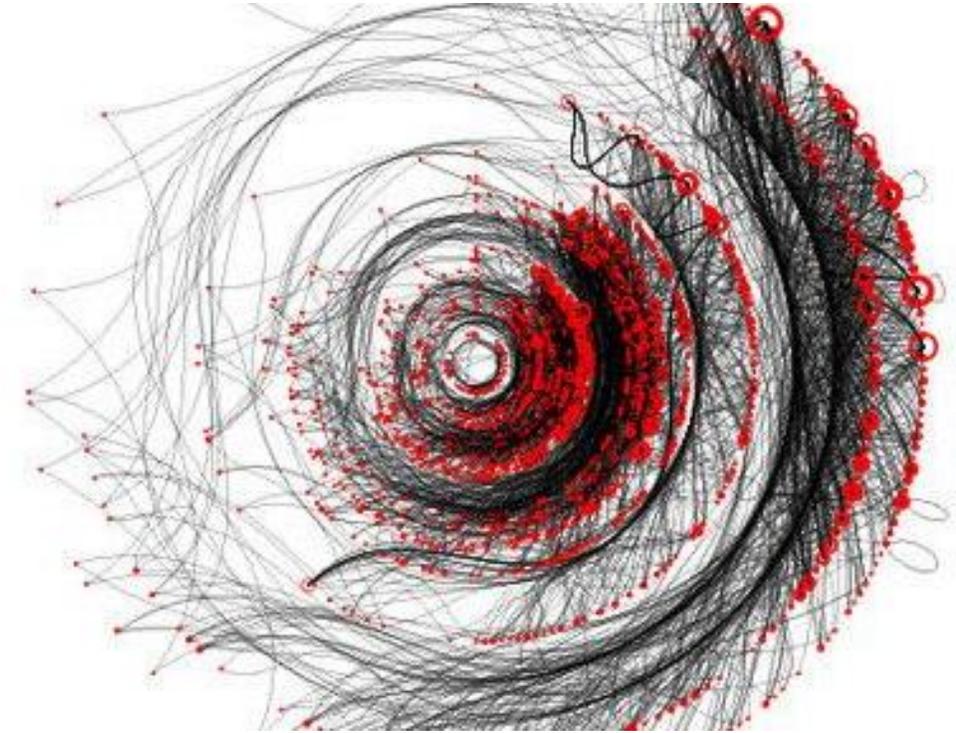
- Data Visualisations help by transforming complex information into something easier to understand.
- Two people can interpret the same data Visualisation differently.
- Visualisations tend to answer "**What**" questions.
- Data storytelling explains the "**Why**"
- It contributes to the Visualisation's **interestingness**
- It makes it **relevant**
- weaves data and Visualisations into a narrative tailored to a specific audience
- convey credibility in the analytical approach
- confidence in the results
- Should give compelling set of **insights** that is **actionable** to the **audience**



- Tableau Dashboards supports Storyboards
- Book reference:  
**Visual Data Storytelling with Tableau, First edition**  
by Lindy Ryan  
Publisher: Addison-Wesley Professional  
*Release Date: May 2018*  
ISBN: 9780134712963

# ► Goal (function )

- The main goal of data Visualisation is communicating information:
  - In a usable meaningful way
  - Fitting
  - Effective
  - Efficient
  - Elegantly
  - Accurately
  - Outcome driven
- It fulfills its objectives only if it **encodes the given input** in such a manner that our **eyes can recognize** and our **brain can comprehend**. Otherwise it is just eye candy or data art.

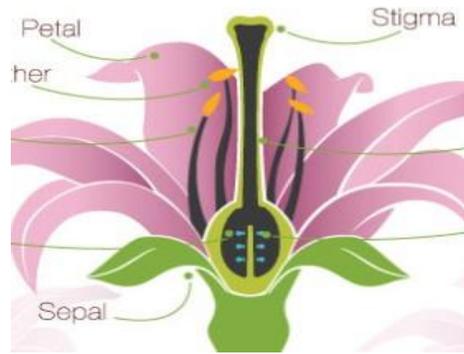


# ► What is involved to “get going”

- **Tableau** – No Programming knowledge required
  - Buy Commercial software – NOT Free
  - Works on MacOS and Windows
  - Much like Pivot tables but much more powerful
  - Uses a WORKBOOK and SHEET file structure, much like Microsoft Excel but with dashboards and storyboards.
- **R and R-Studio**
  - R is a language and environment for statistical computing and graphics
  - R is Open source and Free with thousands of packages to do anything really
  - Download R and R-Studio, Lots of help on the Internet to learn it.
  - R-Studio is a tool that makes programming and visualizing in R easy.
    - R Graphics Plots
    - R Shiny (Web)
      - Shiny is an open source R package that provides an elegant and powerful web framework for building web applications using R. Shiny helps you turn your analyses into interactive web applications without requiring website programming knowledge
    - R Markdown - Reproducible Interactive Research Documents

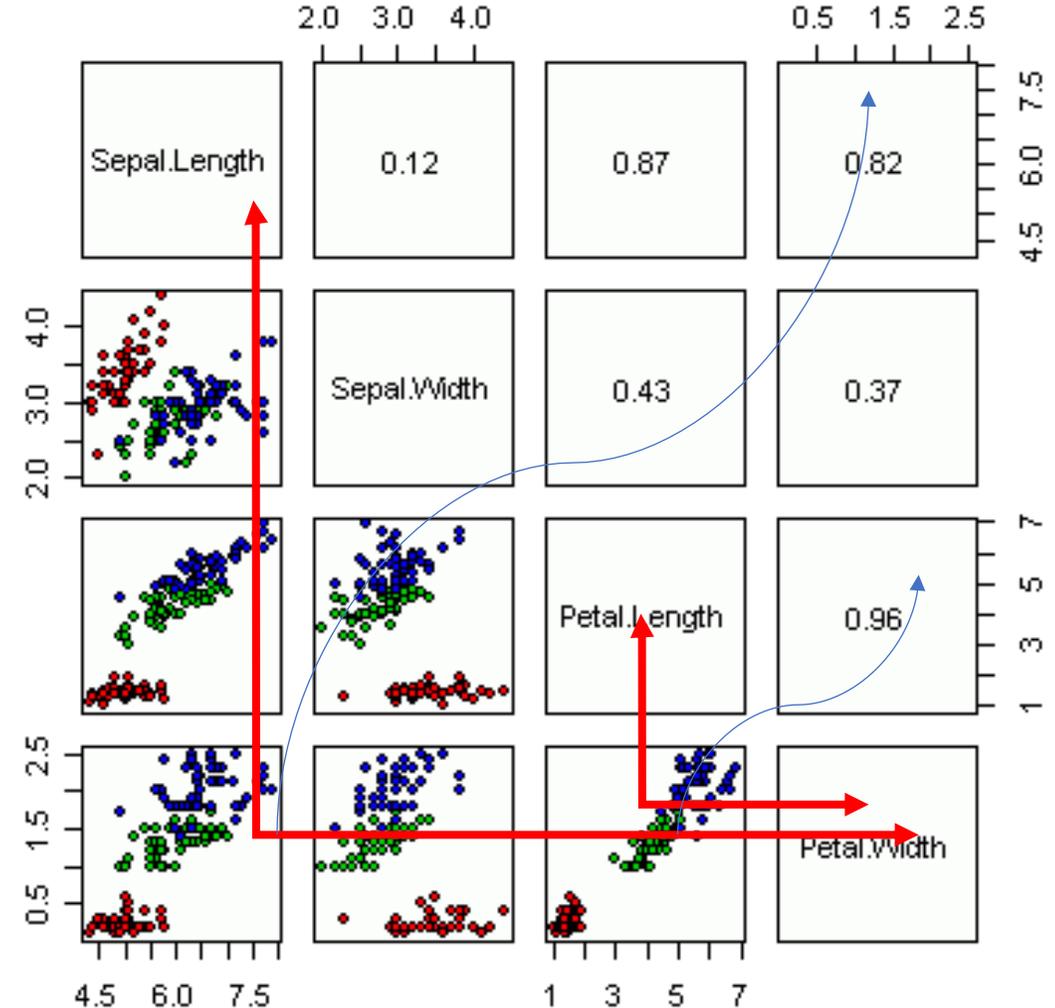
# R Plots: Exploratory Example

```
> iris
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1           5.1         3.5         1.4         0.2  setosa
2           4.9         3.0         1.4         0.2  setosa
3           4.7         3.2         1.3         0.2  setosa
4           4.6         3.1         1.5         0.2  setosa
...
149          6.2         3.4         5.4         2.3 virginica
150          5.9         3.0         5.1         1.8 virginica
```



- This IRIS example data set is built into R
- 50 samples from each of three species of *Iris* flowers
  - *Setosa*
  - *Virginica*
  - *versicolor*
- Four features were measured from each sample: the **length** and the **width** of the **sepals** and **petals**, in centimeters.
- This plot in R allows you to do **Visualise multidimensional scatterplots in a single plot.**
- The Red, Green and Blue indicate the flower species
- As we can see there are a **Linear Relationship** between the Petal Length and Petal Width, with a Pearson Correlation Coefficient of 0.96

## Edgar Anderson's Iris Data



```
> pairs(iris[1:4], main = "Edgar Anderson's Iris Data", pch = 21,
bg = c("red", "green3", "blue")[unclass(iris$Species)],
upper.panel=panel.pearson)
```

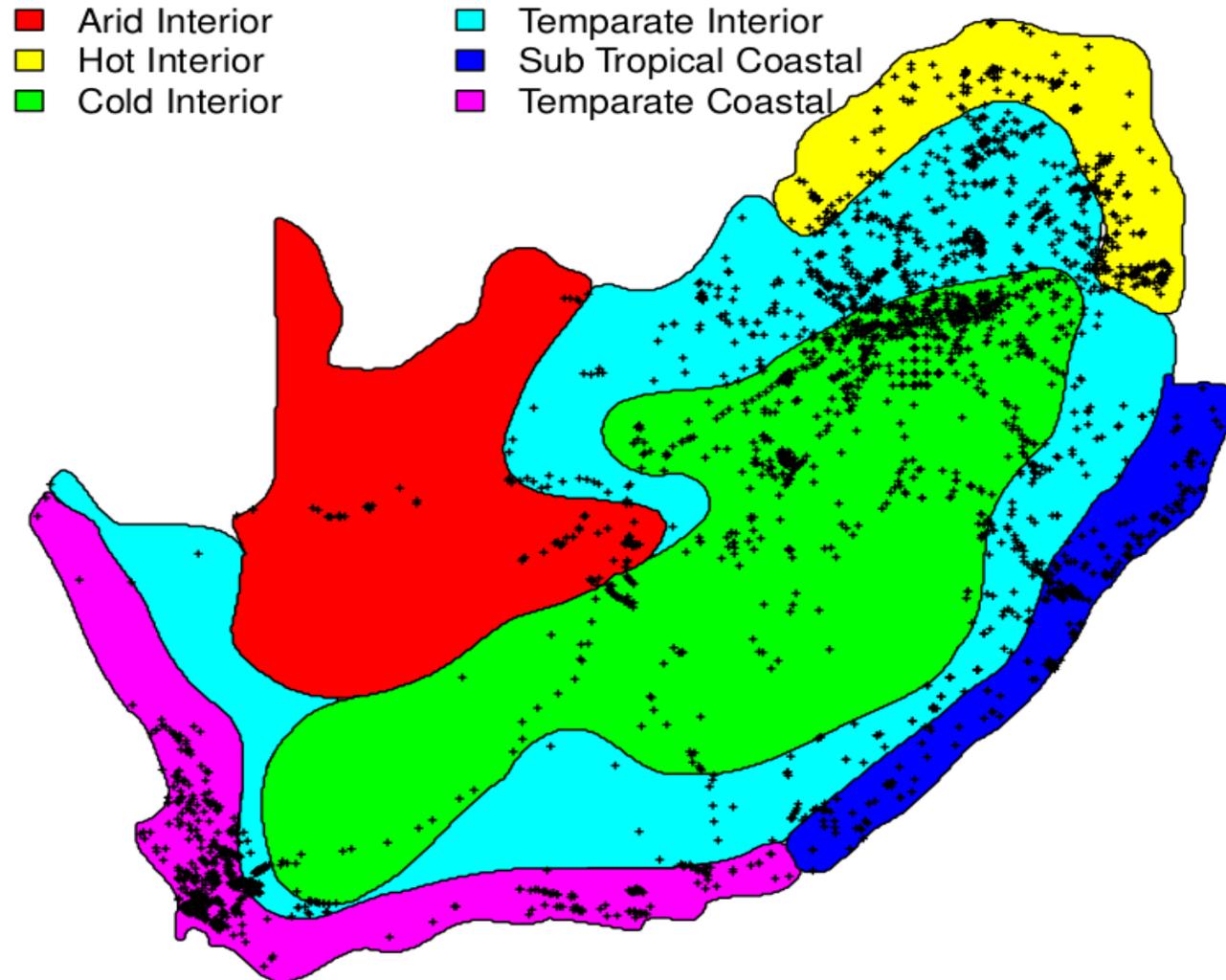
# Regional Load profile classes Visualisation Example

- Objective: Derive regional variants for existing national industrial and commercial customer load profile subclasses, used in Eskom distribution network planning.
- The question we addressed was:  
    **“How do subclass load profiles appear to behave differently in different areas”?**

..In theory, differences in regional behavior are largely due to differences in regional drivers.. and some chance.

- Regional spatial segmentation systems were tested for load profile significance, finding possible spatially linked drivers.
- Final objective: Choose a regional spatial layer that mostly expresses significant real variation in a subclass, per subclass.

# Regional spatial segmentation systems: Climatic zones, R Plot

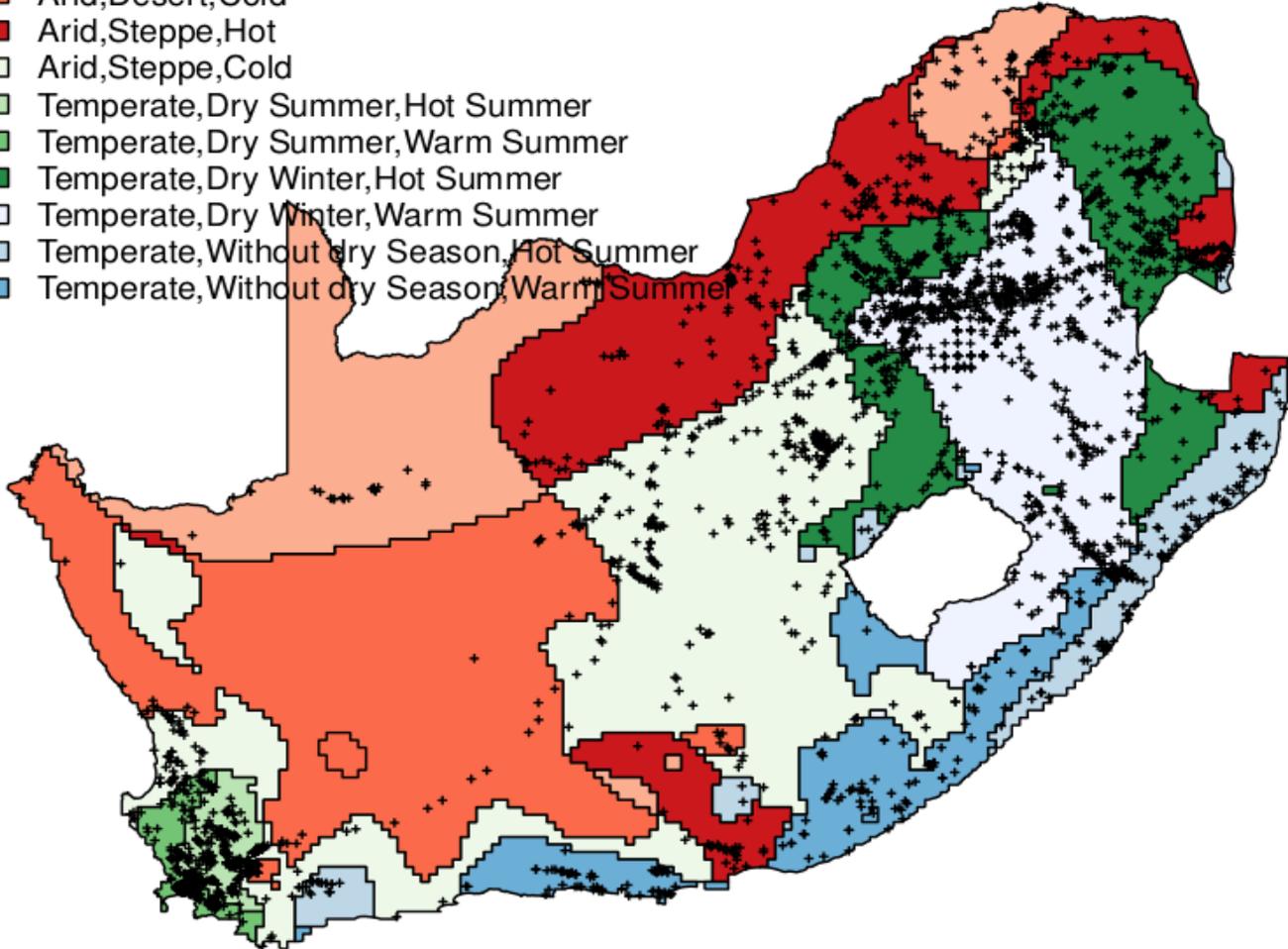


Eskom LPU customers on 30 minute interval metering with their GPS locations

- SANS 204-2 map of South Africa based on six climatic zones.

# Regional spatial segmentation systems: Köppen-Geiger Climatic Zone , R Plot

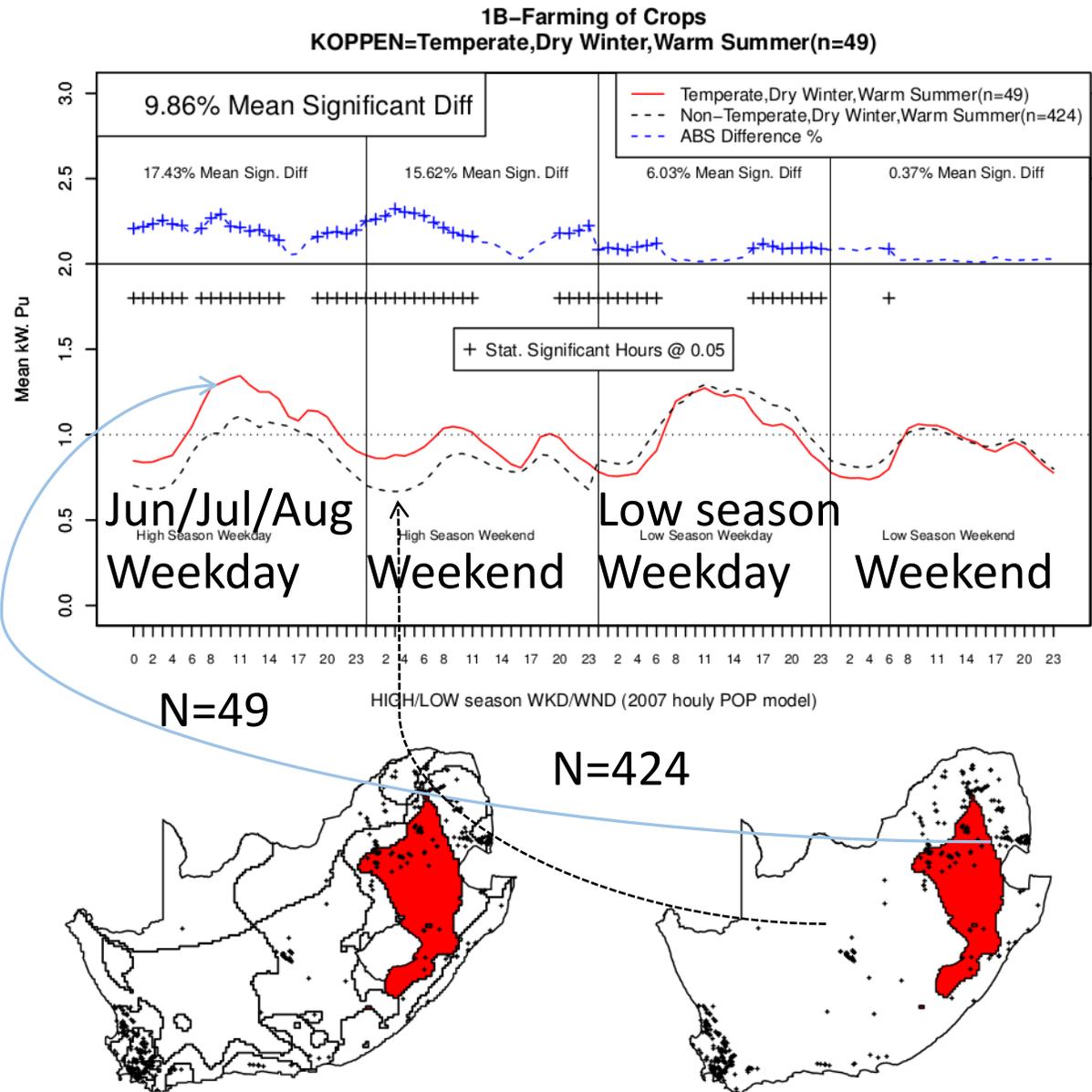
- Tropical,Savannah,
- Arid,Desert,Hot
- Arid,Desert,Cold
- Arid,Steppe,Hot
- Arid,Steppe,Cold
- Temperate,Dry Summer,Hot Summer
- Temperate,Dry Summer,Warm Summer
- Temperate,Dry Winter,Hot Summer
- Temperate,Dry Winter,Warm Summer
- Temperate,Without dry Season,Hot Summer
- Temperate,Without dry Season,Warm Summer



Eskom LPU customers on 30 minute interval metering with their GPS locations

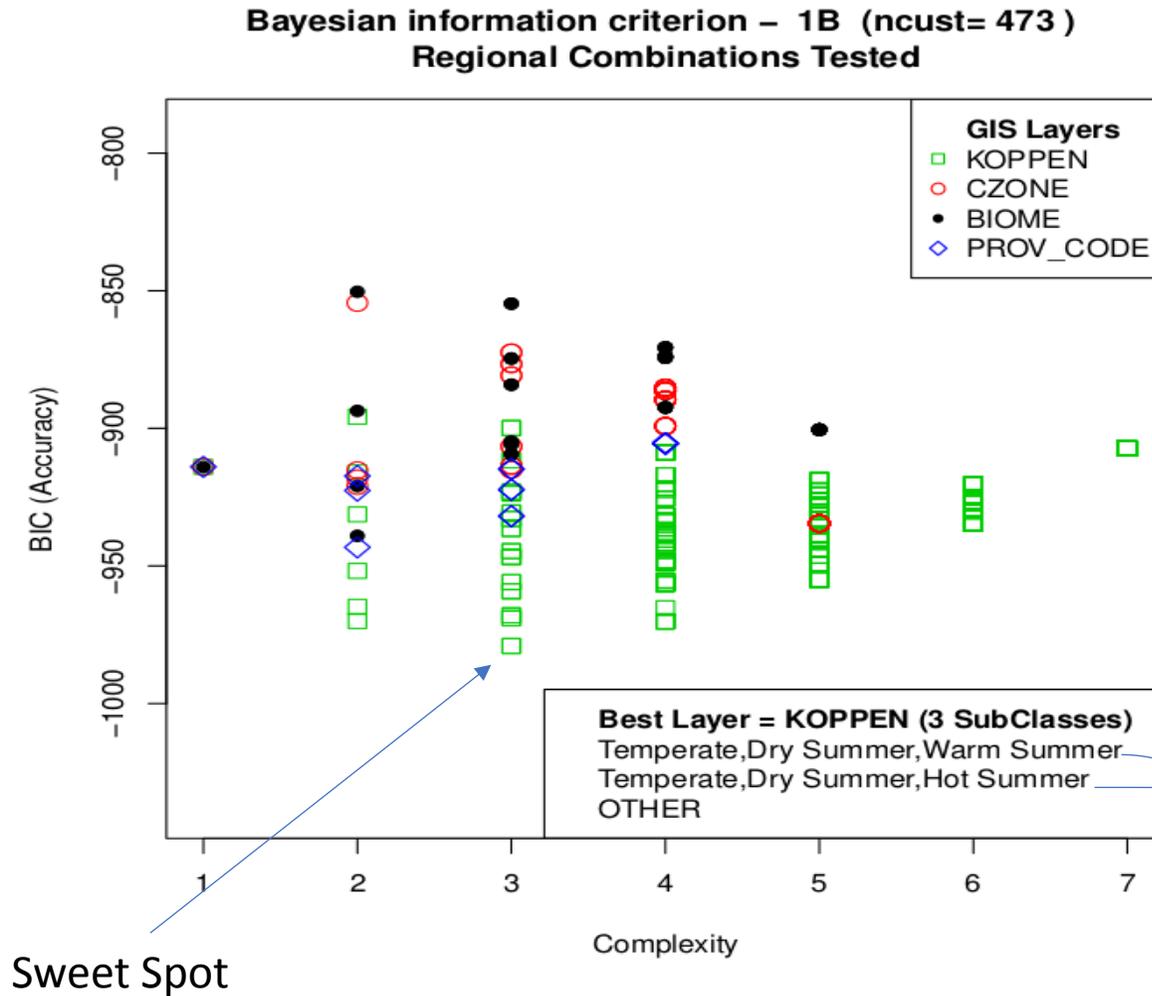
- Köppen-Geiger most widely used climatic classification system
- Based on climatic system used by ancient Greeks and modified by plant specialists
- Combining vegetation temperature and rainfall

# Regional Load profile classes plotted in R



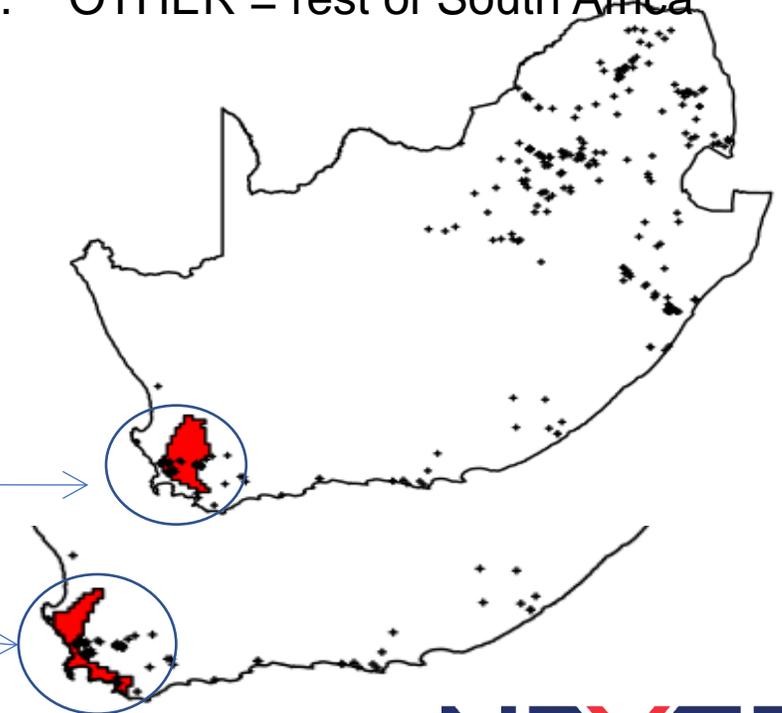
- Multiple spatial segmentation systems and layers were tested
- Dividing South Africa into sub areas for all possible combinations and subclass complexity levels, per economic sector
- Done for all and spatial systems and layers as seen on this picture.
- Average daily energy normalised load profiles per season and day type was computed for each 2 areas over a year.
- In this example the solid red line is from the Köppen-Geiger Climatic zone system layer: **“Temperate-Dry Winter-Warm Summer”** tested against the **Rest of South Africa** not in that region as the dashed black line.

# Optimal Regional Load profile classes plotted in R



These Regional Subclasses are optimal using Köppen Geiger system increasing accuracy while limiting complexity:

1. Temperate Dry, Hot, Summer
2. Temperate Dry, Warm, Summer
3. OTHER = rest of South Africa



# R Shiny: Market Intelligence Tool

Market Intelligence Master Data Management Correlation Matrix SA Media Data Collection

MTS Energy per Regions

- MTS Stations per Region
  - CENTRAL REGION
  - INTERNATIONAL
  - NORTHWEST
  - WESTERN REGION
  - EASTERN REGION
  - WESTERN REGION
    - ACORNHOEK
    - ARARAT
    - BIGHORN
    - DINALEDI
    - FOSKOR
    - HENDRINA
    - KOMATI
    - KOMATIPOORT

SA Media Full Text Search for: eskom

Search  Geosearch (within 100km of selected point)

Full Text search through all Articles.

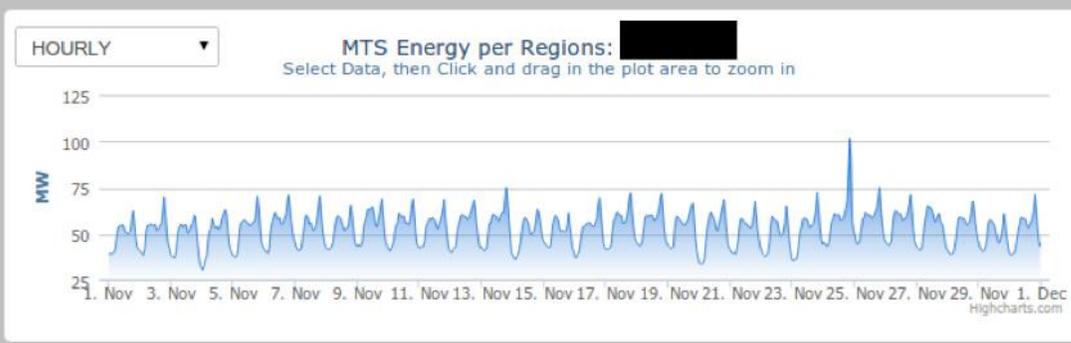
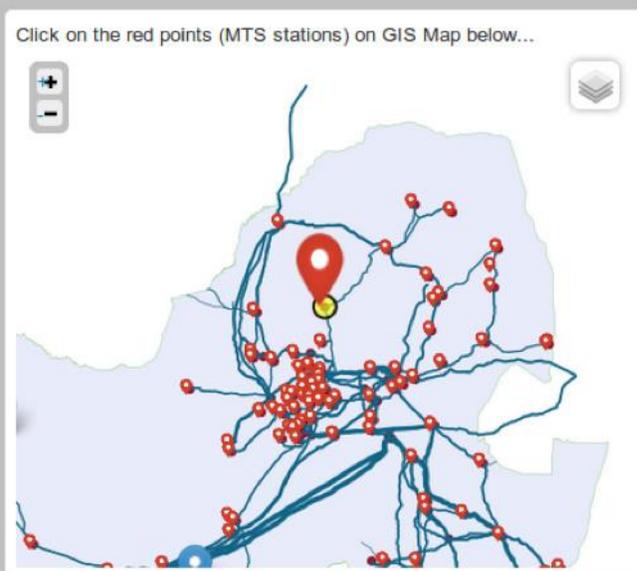
Top 100 Search Results: SA Media Articles (from local MongoDB document store)

id	title	date	topic	text_score	senti	senti_score	subj_score
3781773	ALLIANCE 'UNCOMFORTABLE' WITH ESKOM HIKES	2009-11-16	ECONOMIC AFFAI	102	pos	0.049237	0.29503
3773610	POWER OF THE JUGGERNAUT	2009-10-26	ECONOMIC AFFAI	102	pos	0.09883	0.45991
3746677	ESKOM WORKERS TO MARCH AS STRIKE LOOMS	2009-08-07	LABOUR (1)	95	pos	0.077325	0.33738
4182059	INGULA ON TRACK, ESKOM SAYS	2013-09-02	ECONOMIC AFFAI	95	pos	0.13468	0.50856
3572578	WESIZWE BUDGETS FOR POWER	2008-04-01	ECONOMIC AFFAI	89	pos	0.18764	0.53846
3702787	EXPECT NO MERCY - NUM	2009-04-20	LABOUR (1)	89	pos	0.068382	0.46389

Page 1 of 11 View 1 - 6 of 66

All geocoded locations for current selected document...

free state (-29.121065,26.214003)  Show document locations on map (blue) [Download selected article as PDF](#)



Add custom polygon GIS Shape File(.shp) + Attributes(.dbf) file: (ESRI map polygons)

No file chosen

0%

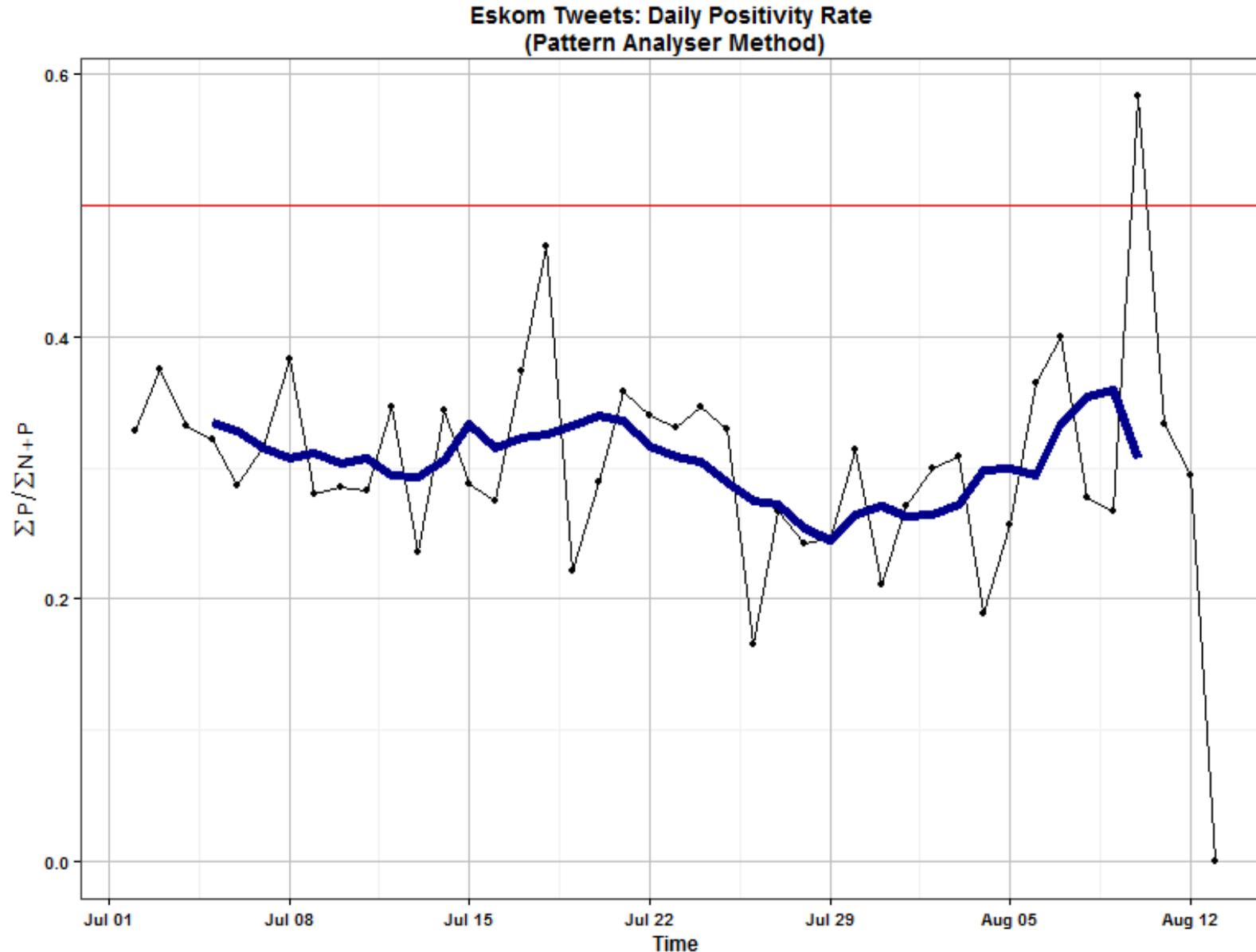
Month: 1 11 12

Year: 2000 2009 2012

- Map and full Text search on South African media and social media articles that has been converted to text, geo-coded, sentiment analyzed and stored in a No-SQL document database.
- The map shows Transmission and distribution grids with MTS stations
- The timeseries shows MTS Energy hourly load profile per week linked to the timestamp of the article.
- The hierarchy shows MTS stations per region.
- The above visualised components and geocoded articles are all interconnected.
- Clicking (filtering) on one dimension updates the other dimensions.
- Geo-Search within selected radius in kilometers or supply area and can be combined with full text search.
- You can also Geo-Search within custom GIS Shape file.
- Read Articles

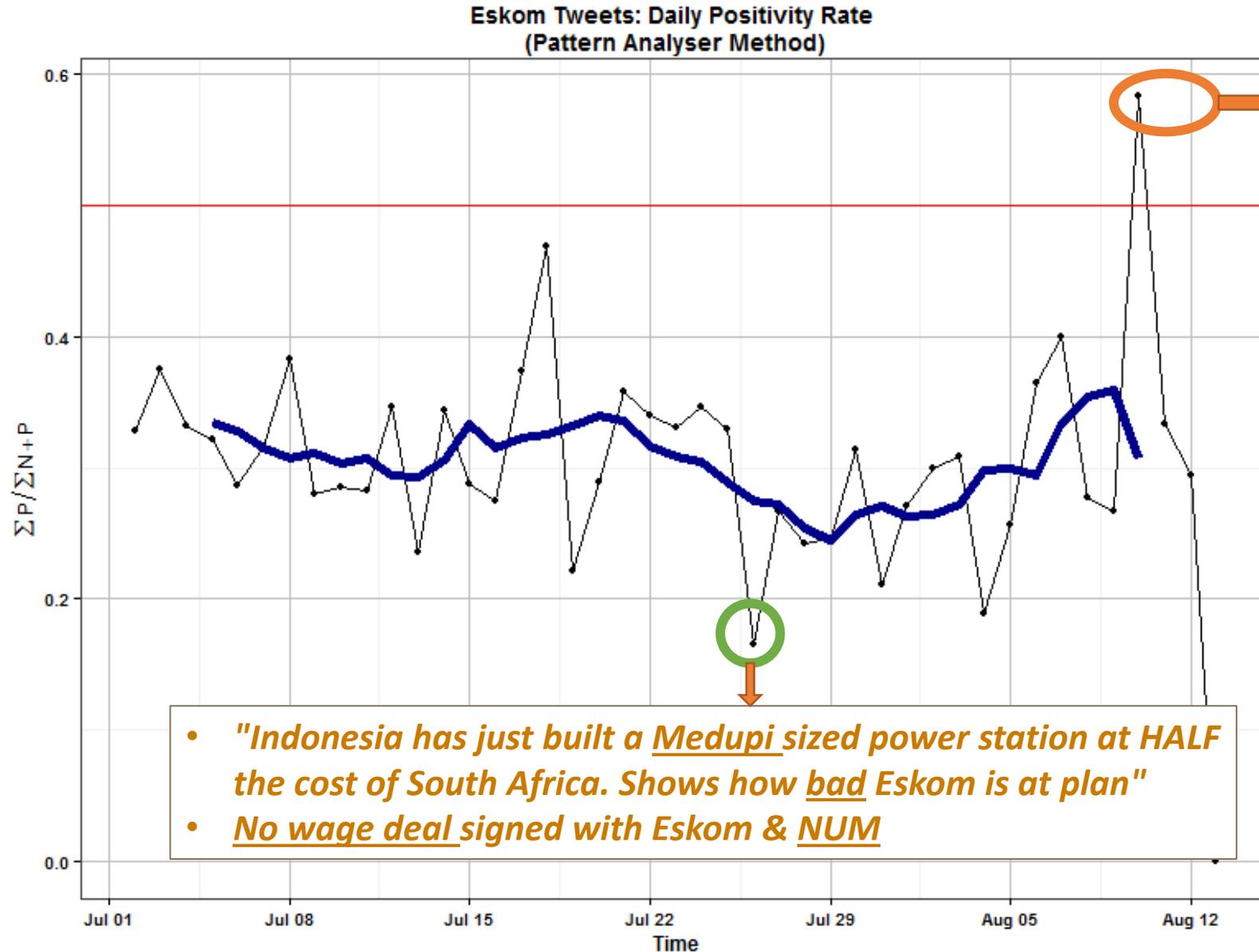


# Overall Daily Sentiment for Tweets containing term Eskom (July, Aug 2013) $n=14254$



- Two Months of Twitter data
- Tweets with Term ESKOM
- Sentiment analysis was applied to the daily tweets. (Positive / Negative)
- This is hard to interpret without a Storyboard

# Overall Daily Sentiment for Tweets containing term Eskom (July, Aug 2013) $n=14254$



*"LOL Eskom is the only company that begs u not to use their product"*

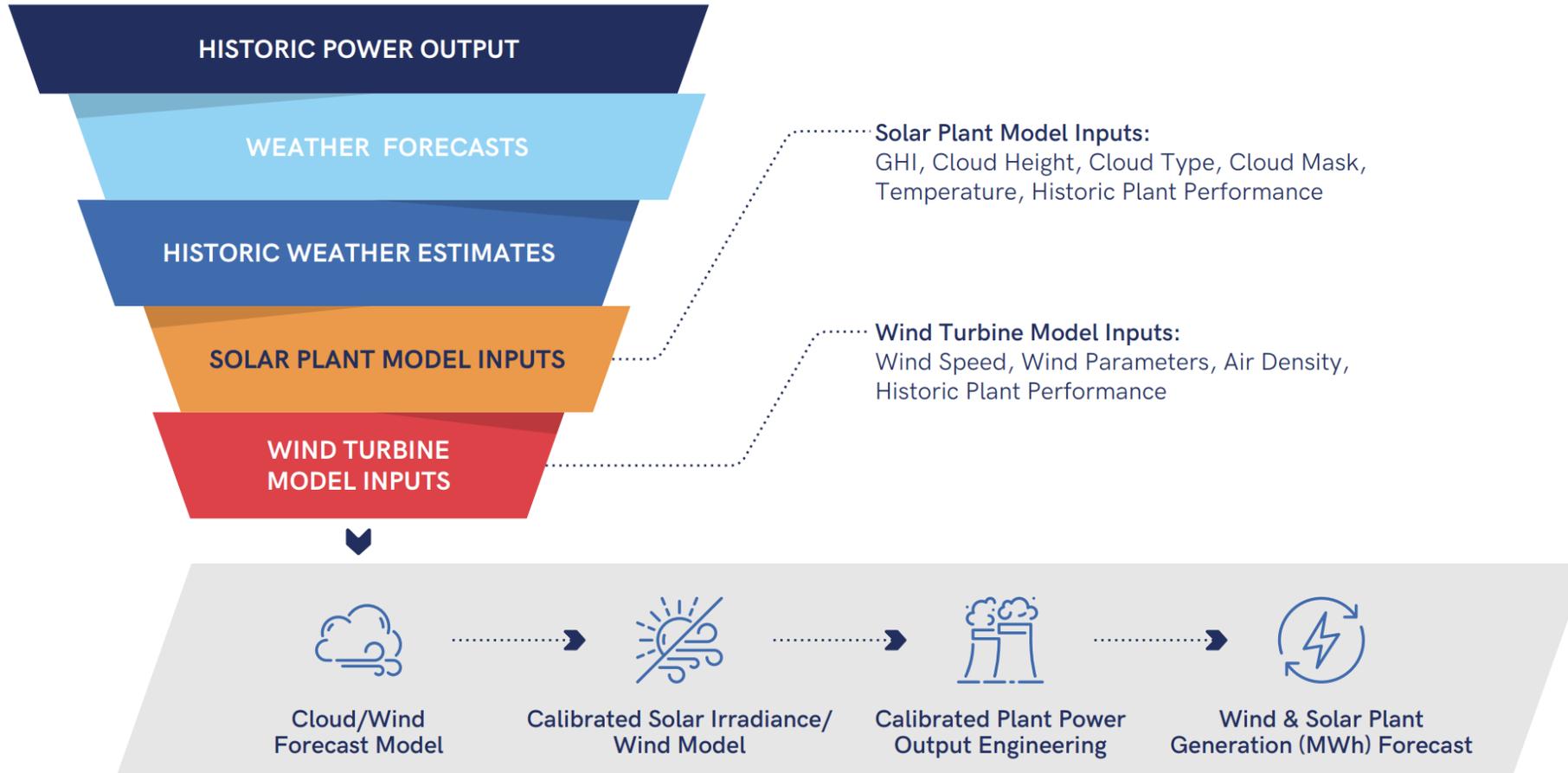
Re-tweeted 64x!

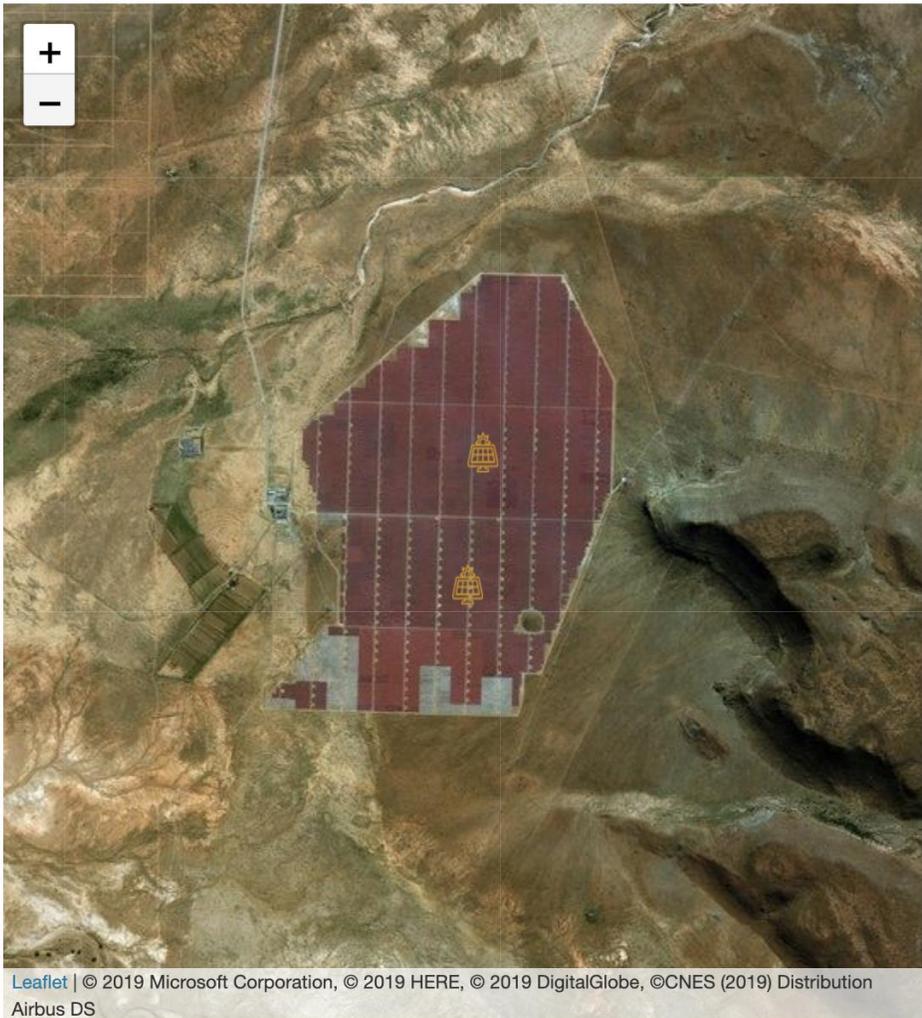
*"LOL" being positive*

Better Visualisation as:  
STORY BOARD

# Renewable Forecasting Visualisation Example in R Shiny

## Renewable Energy Forecasting Model





Show 4 entries

Search:

id	sitename	technology	MW	status
1	Solar Capital De Aar 3	Solar Photovoltaic (PV)	75	Fully operational
2	SlimSun Swartland Solar Park	Solar Photovoltaic (PV)	5	Fully operational
3	RustMo1 Solar Farm	Solar Photovoltaic (PV)	6.8	Fully operational
4	Mulilo Renewable Energy Solar PV De Aar	Solar Photovoltaic (PV)	9.7	Fully operational

Showing 1 to 4 of 104 entries

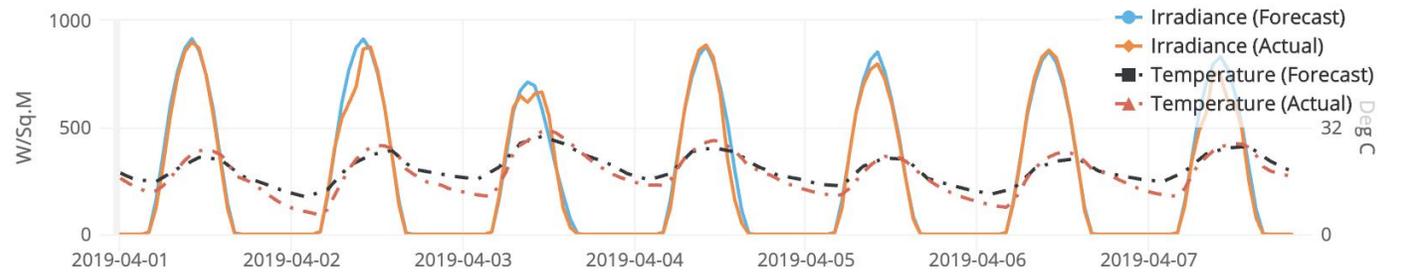
First Previous 1 2 3 4 5 ... 26 Next Last

Date range input: yyyy-mm-dd

2019-04-01 to 2019-04-07

### Historic Actual vs Forecast - Solar Irradiation / Surface Temperature

Solar Capital De Aar 3: |Mean Abs Error = 32.6g W/Sq.M|  
2019-04-01 to 2019-04-07



Historic Weather Forecast Error

3 Days Ahead (Solar/Temperature) Forecast

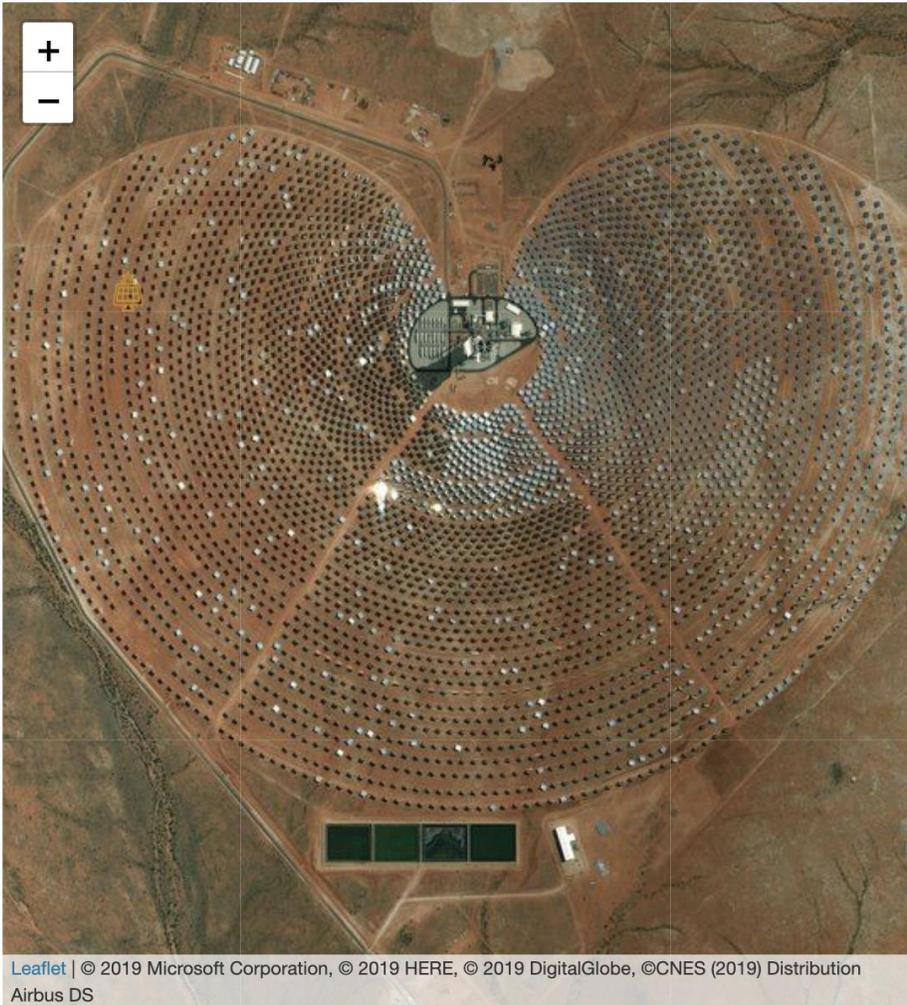
127.0.0.1:7637/#

IPP Projects

Solar PV Model

Wind Model

Our models use advanced weather models and satellite data for forecasts in R. We use R Shiny as the Visualisation tool. This screen show a spatial temporal view per IPP site.



Show 4 entries

Search:

id	sitename	technology	MW	status
28	Khi Solar One	Concentrated Solar Thermal (CSP)	50	Construction
29	KaXu Solar One	Concentrated Solar Thermal (CSP)	100	Fully operational
45	Bokpoort CSP Project	Concentrated Solar Thermal (CSP)	50	Construction
63	Xina CSP South Africa	Concentrated Solar Thermal (CSP)	100	Awaiting construction (approved & financed)

Showing 1 to 4 of 7 entries (filtered from 104 total entries)

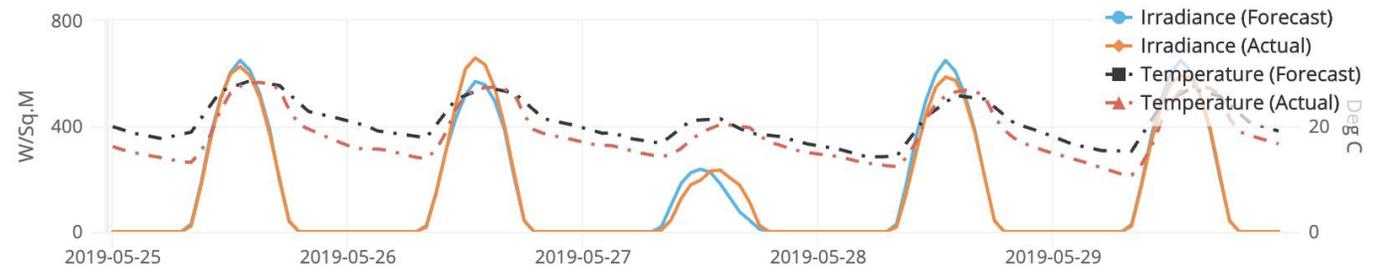
First Previous **1** 2 Next Last

Date range input: yyyy-mm-dd

2019-05-25 to 2019-05-29

### Historic Actual vs Forecast - Solar Irradiation / Surface Temperature

Khi Solar One: |Mean Abs Error = 29,32 W/Sq.M|  
2019-05-25 to 2019-05-29



Historic Weather Forecast Error

3 Days Ahead (Solar/Temperature) Forecast

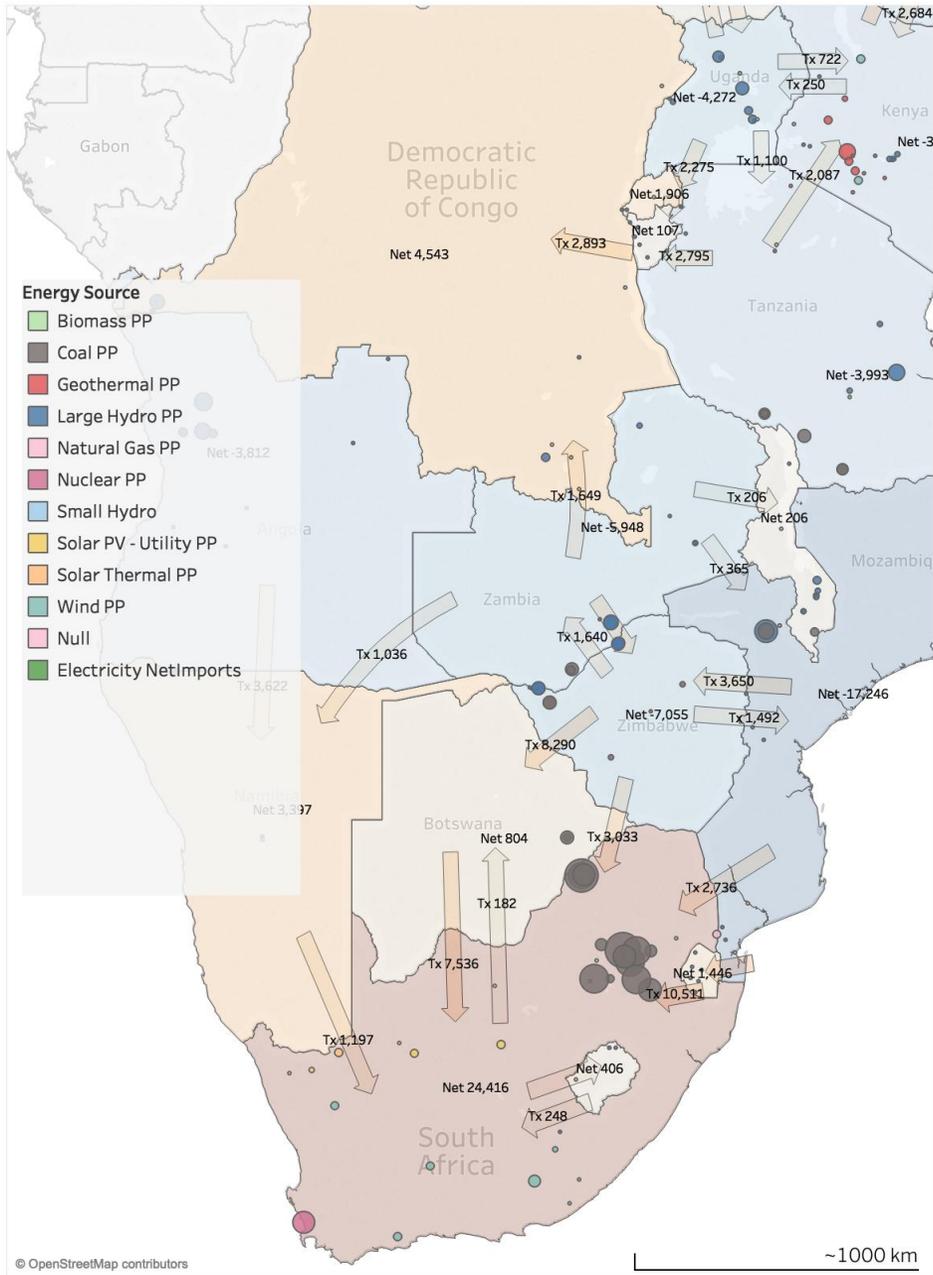
IPP Projects

Solar PV Model

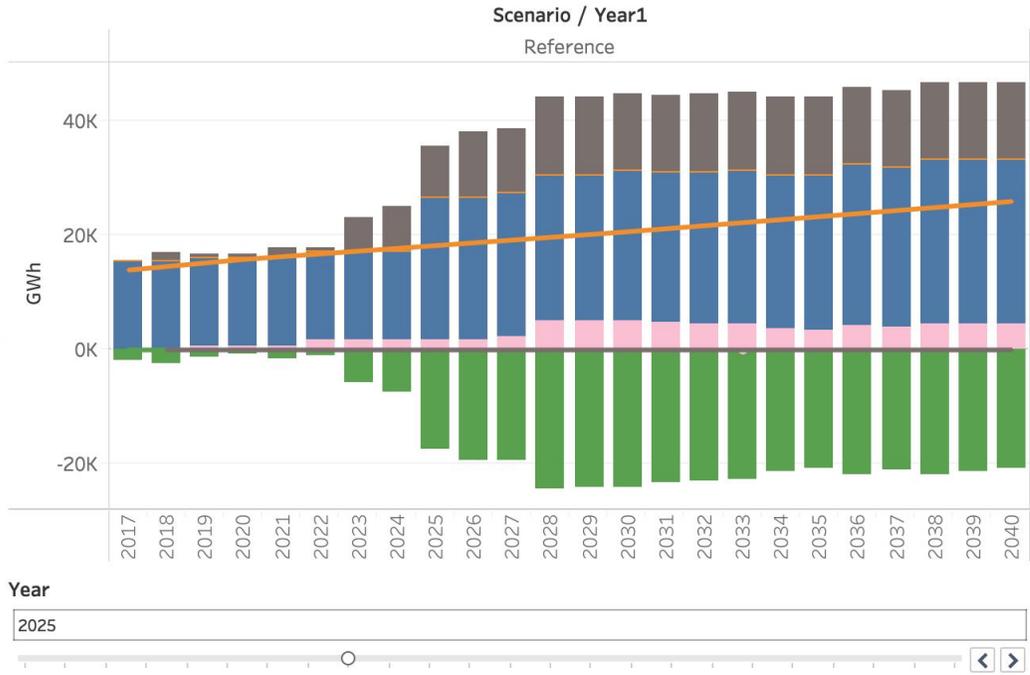
Wind Model

Another example of a CSP plant along with it's historic Irradiation forecast performance, spatial temporal view in Shiny

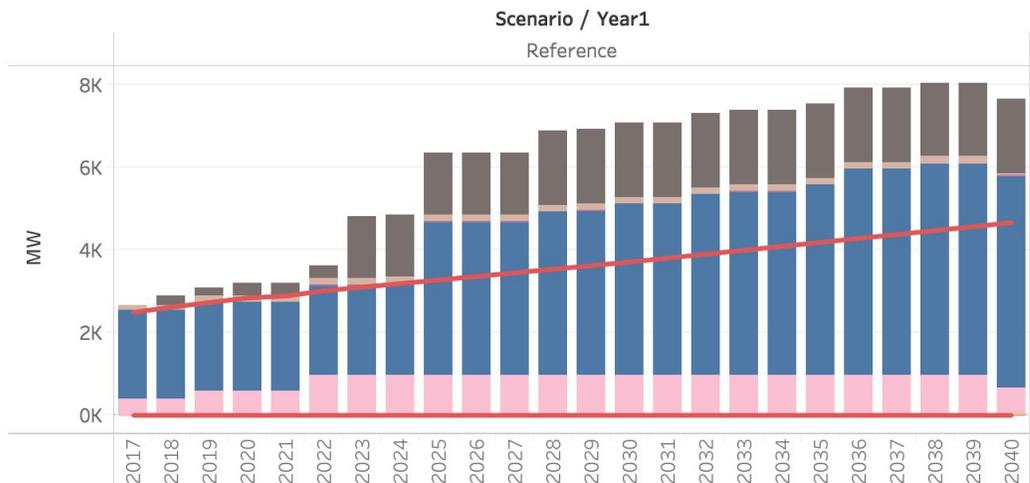




### Generation Output (All) Mozambique



### Total Installed Capacity (All) Mozambique



WithZK

- In
- Out

Scenario

- Delayed Interconnectors
- Dry Climate
- Pessimistic
- Reference

Country Name

- (All)
- Angola
- Botswana
- Burundi
- DRC
- Egypt
- Ethiopia
- Kenya
- Lesotho
- Malawi
- Mozambique
- Namibia
- Rwanda
- South Africa
- South Sudan
- Sudan
- Swaziland
- Tanzania
- Uganda

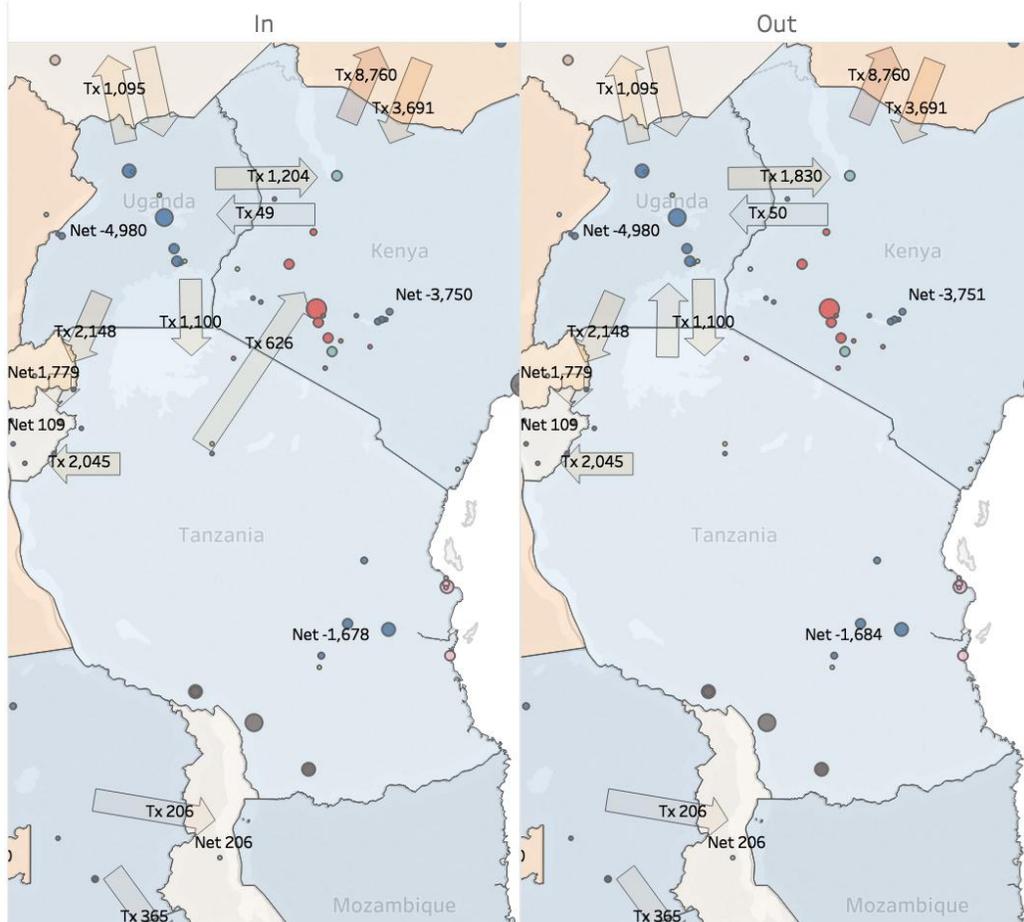
# Example Storyboard: Impact of ZTK line being build on Electricity Trading in 2024

In 2024 on the Reference Scenario the Change in Electricity Trade between Tanzania and Kenya will increase by 626 GWh as a consequence of the ZTK interconnector being build.

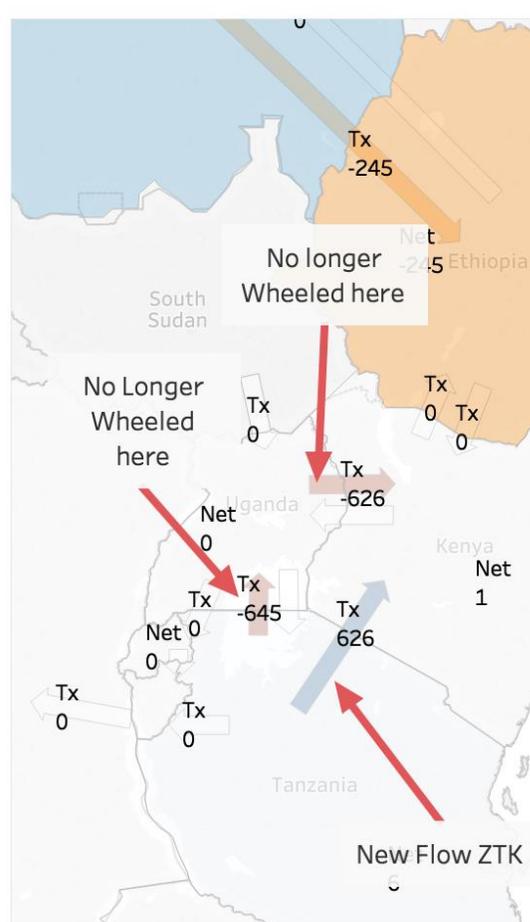
In 2024 zooming out and selecting the Dry Climate Scenario reveals that in the scenario the Trade between Zambia and Namibia will increase with some energy being wheeled into South Africa as a consequence of the ZTK interconnector.

## SAPP/EAPP - Compare - 2024 Reference

WithZTK



## SAPP/EAPP (IN) Change Map - 2024 Reference



Year



Scenario

- Delayed Interconne...
- Dry Climate
- Pessimistic
- Reference

Energy Source

- Biomass PP
- Coal PP
- Diesel PP
- Geothermal PP
- HFO PP
- Large Hydro PP
- Natural Gas PP
- Nuclear PP
- Small Hydro
- Solar PV - Utility PP
- Solar Thermal PP
- Wind PP
- Null
- Electricity NetImports

**NEXTEC**

# ► Potential Use Cases

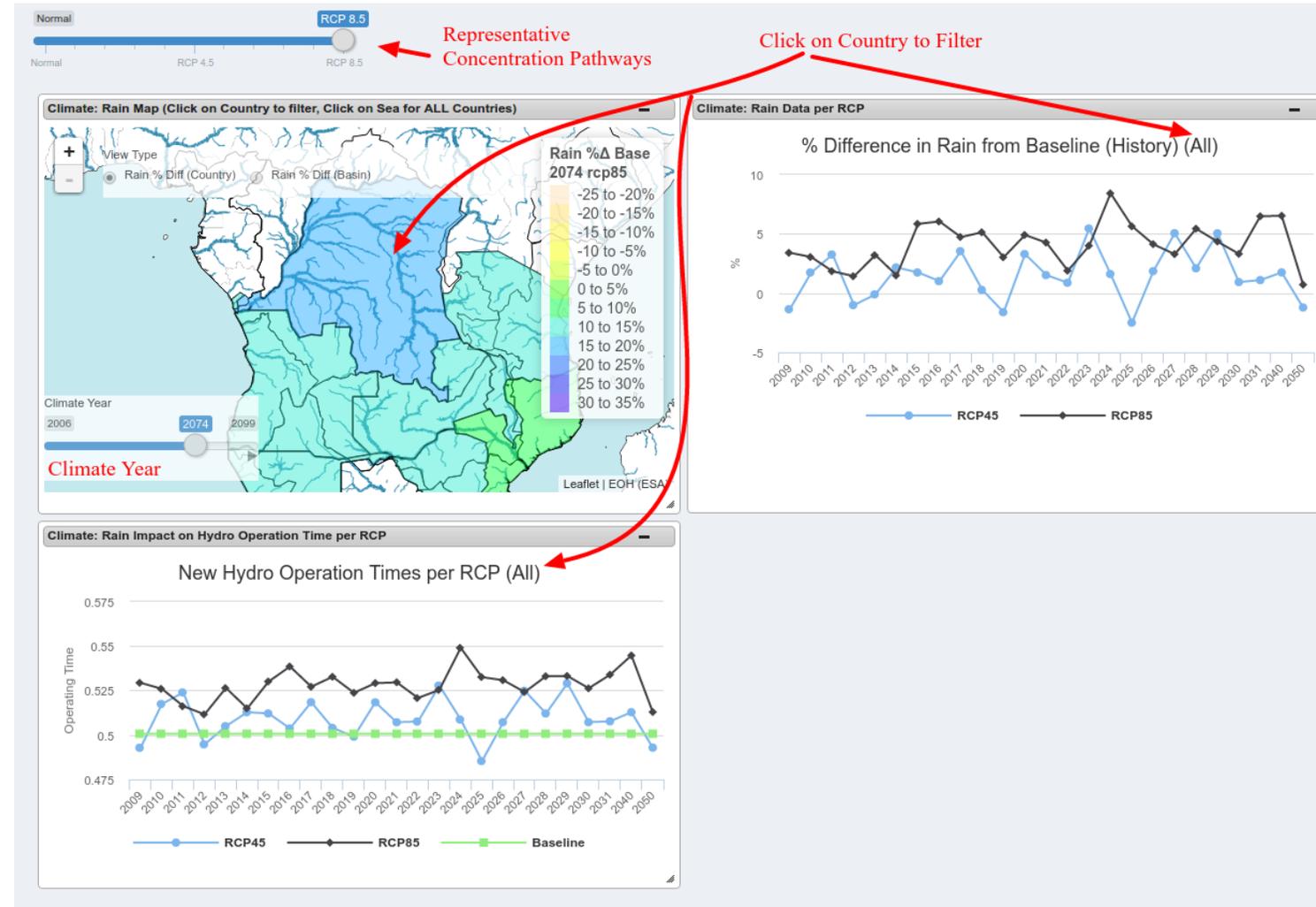
Testing the impacts on SAPP for various scenarios from medium term to near term/operational

Such as:

- 1) Impact of new TX Projects/Interconnectors
- 2) New Generation
- 3) Fuel Prices
- 4) Outage Analysis
- 5) Impact of Renewables
- 6) Average and Marginal Cost of Gx for Country/Pool under various scenarios
- 7) CO2 Emissions

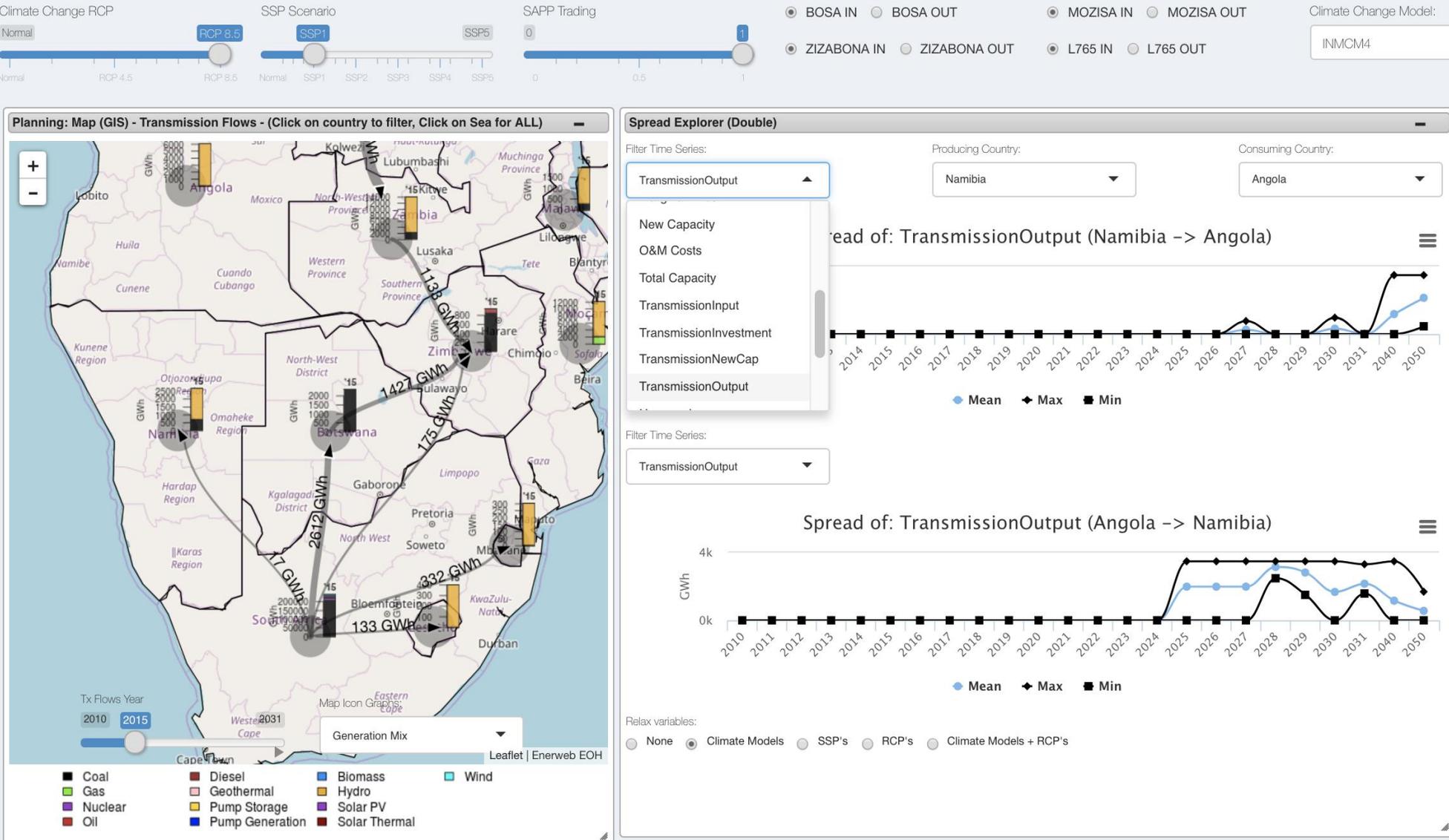
# Climate Resilient Infrastructure Development Facility (CRIDF)

- R Shiny was used to Visualise this model.
- Impact of Climate change on Electricity Trade
- Energy Model Based on IRENA SPLAT
- Average weekly temperature and rainfall data per country and Representative Concentration Pathway (RCP)
- NASA Earth Exchange Global Daily Downscaled Projections model ensemble from the 21 climate models, using R on Amazon Web Services.
- The raw dataset of these climate projections is downscaled at a spatial resolution of approximately 25 km x 25 km)
- Data archive size of more than 12 TB (1 Terabyte = 1012 Megabytes) on the Amazon Cloud.



# Climate Resilient Infrastructure Development Facility (CRIDF)

- Over a 1000 least cost models results where produced using high performance computing.
- Representing all the combinations of the 21 Climate change models
- RCP slider
- Interconnectors ON/OFF
- And other variables
- A Spread explorer was developed to see the range of possible interconnector flows or other variable if one variable is **relaxed** for **sensitivity analysis**



# ► Getting started LR Data Visualisation Web Links

Links with examples for self learning:

- **R and R-Studio**

- R (<https://cran.r-project.org/bin/windows/base/> , <https://www.rstudio.com/>)
- R Plots (\* <https://www.r-graph-gallery.com/> )
- R Shiny (\* <https://shiny.rstudio.com/gallery/> )
- R Markdown (\* <https://rmarkdown.rstudio.com/> )

- **Tableau** (<https://www.tableau.com/> )

- <https://public.tableau.com/en-us/s/gallery/>
- <https://public.tableau.com/en-us/s/resources>



END

▶ Questions?

<http://nextec-esa.co.za>