Data Visualization

Final Project

Date: 24.01.22

Professor: Gilad Ravid

Student: Dafna Meron

Contents

Introduction

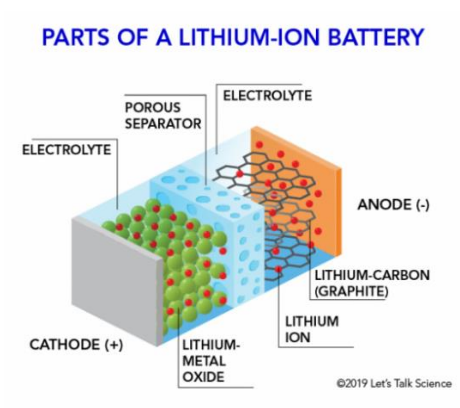
The Organization

StoreDot is an Israeli startup company based in Herzliya. The company develops extreme fast charged (XFC) Lithium-ion batteries based on materials innovation, design and synthesis of organic & inorganic compounds. The company's vision relates to the world of electrical vehicles (EV) that currently faces many challenges: low driving range, high battery cost, high battery weight, and mainly long recharging time - which can take several hours. StoreDot's innovative technology addresses all these issues and mainly aims to turn a 400 KM recharging experience into 5 minutes.

The development process focuses on finding new materials and configurations for a battery, which will provide the required performance according to the company's vision, with maximum safety, and minimum damage to natural resources. StoreDot develops all the components that make a battery cell: Anode, Cathode & Electrolyte. The company also develops the entire system that eventually integrates all these components into a full battery pack with dozens of cells.

Short explanation on the basic cell

Lithium-ion battery is a type of rechargeable battery. Lithium ions move from the **negative electrode (anode)** through an **electrolyte** to the **positive electrode (cathode)** during discharge, and back when charging.

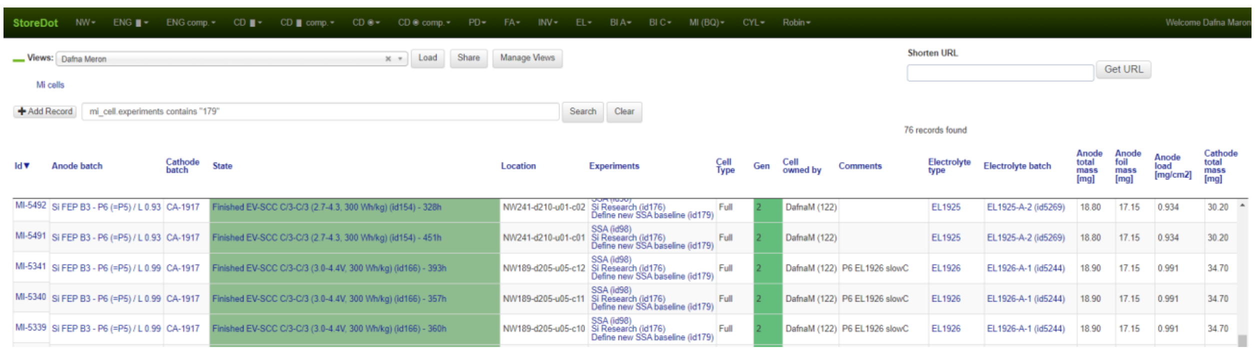


Data

Data Source

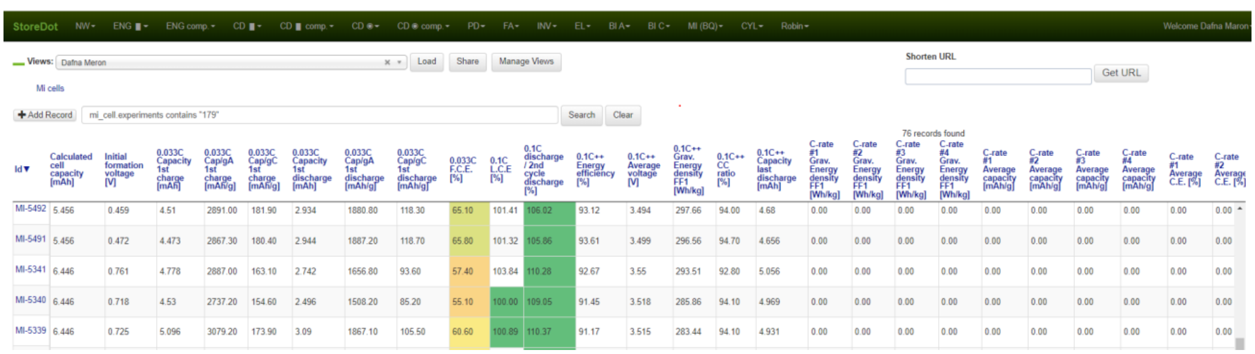
The data source is based on StorDot's original database. The main table is 'Cells' – each entity describes an experiment of a different battery cell configuration. The first columns hold the experiment conditions – set parameters, and the later columns hold the experiment results – measured parameters. Other tables contain data on the materials of each cell component, product types, cell types, cell owners, experiment IDs and more. This data source holds hundreds of thousands of entries dating back to 2017.

Example of experiment conditions – set parameters:



Id	Anode batch	Cathode batch	State	Location	Experiments	Cell Type	Gen	Cell owned by	Comments	Electrolyte type	Electrolyte batch	Anode total mass [mg]	Anode foil mass [mg]	Anode load [mg/cm2]	Cathode total mass [mg]
MI-5492	Si FEP B3 - P6 (=P5) / L 0.93	CA-1917	Finished EV-SCC C/3-C/3 (2.7-4.3, 300 Wh/kg) (id154) - 328h	NW241-d210-u01-c02	Si Research (id176) Define new SSA baseline (id179)	Full	2	DafnaM (122)		EL1925	EL1925-A-2 (id5269)	18.80	17.15	0.934	30.20
MI-5491	Si FEP B3 - P6 (=P5) / L 0.93	CA-1917	Finished EV-SCC C/3-C/3 (2.7-4.3, 300 Wh/kg) (id154) - 451h	NW241-d210-u01-c01	SSA (id98) Si Research (id176) Define new SSA baseline (id179)	Full	2	DafnaM (122)		EL1925	EL1925-A-2 (id5269)	18.80	17.15	0.934	30.20
MI-5341	Si FEP B3 - P6 (=P5) / L 0.99	CA-1917	Finished EV-SCC C/3-C/3 (3.0-4.4V, 300 Wh/kg) (id166) - 393h	NW189-d205-u05-c12	SSA (id98) Si Research (id176) Define new SSA baseline (id179)	Full	2	DafnaM (122)	P6 EL1926 slowC	EL1926	EL1926-A-1 (id5244)	18.90	17.15	0.991	34.70
MI-5340	Si FEP B3 - P6 (=P5) / L 0.99	CA-1917	Finished EV-SCC C/3-C/3 (3.0-4.4V, 300 Wh/kg) (id166) - 357h	NW189-d205-u05-c11	SSA (id98) Si Research (id176) Define new SSA baseline (id179)	Full	2	DafnaM (122)	P6 EL1926 slowC	EL1926	EL1926-A-1 (id5244)	18.90	17.15	0.991	34.70
MI-5339	Si FEP B3 - P6 (=P5) / L 0.99	CA-1917	Finished EV-SCC C/3-C/3 (3.0-4.4V, 300 Wh/kg) (id166) - 360h	NW189-d205-u05-c10	SSA (id98) Si Research (id176) Define new SSA baseline (id179)	Full	2	DafnaM (122)	P6 EL1926 slowC	EL1926	EL1926-A-1 (id5244)	18.90	17.15	0.991	34.70

Example of experiment results – measured parameters:



Id	Calculated cell capacity [mAh]	Initial formation voltage [V]	0.033C Capacity 1st charge [mAh]	0.033C Cap/gA 1st charge [mAh/g]	0.033C Cap/gC 1st charge [mAh/g]	0.033C Capacity 1st discharge [mAh]	0.033C Cap/gA 1st discharge [mAh/g]	0.033C Cap/gC 1st discharge [mAh/g]	0.033C F.C.E. [%]	0.1C L.C.E [%]	0.1C discharge / 2nd cycle discharge [%]	0.1C++ Energy efficiency [%]
MI-5492	5.456	0.459	4.51	2891.00	181.90	2.934	1880.80	118.30	65.10	101.41	106.02	93.12
MI-5491	5.456	0.472	4.473	2867.30	180.40	2.944	1887.20	118.70	65.80	101.32	105.86	93.61
MI-5341	6.446	0.761	4.778	2887.00	163.10	2.742	1656.80	93.60	57.40	103.84	110.28	92.67
MI-5340	6.446	0.718	4.53	2737.20	154.60	2.496	1508.20	85.20	55.10	100.00	109.05	91.45
MI-5339	6.446	0.725	5.096	3079.20	173.90	3.09	1867.10	105.50	60.60	100.89	110.37	91.17

Id	0.1C++ Average voltage [V]	0.1C++ Grav. Energy density FF1 [Wh/kg]	0.1C++ CC ratio [%]	0.1C++ Capacity last discharge [mAh]	C-rate #1 Grav. Energy density FF1 [Wh/kg]	C-rate #2 Grav. Energy density FF1 [Wh/kg]	C-rate #3 Grav. Energy density FF1 [Wh/kg]	C-rate #4 Grav. Energy density FF1 [Wh/kg]	C-rate #1 Average capacity [mAh/g]	C-rate #2 Average capacity [mAh/g]	C-rate #3 Average capacity [mAh/g]	C-rate #4 Average capacity [mAh/g]	C-rate #1 Average C.E. [%]	C-rate #2 Average C.E. [%]
MI-5492	3.494	297.66	94.00	4.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MI-5491	3.499	296.56	94.70	4.656	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MI-5341	3.55	293.51	92.80	5.056	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MI-5340	3.518	285.86	94.10	4.969	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MI-5339	3.515	283.44	94.10	4.931	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

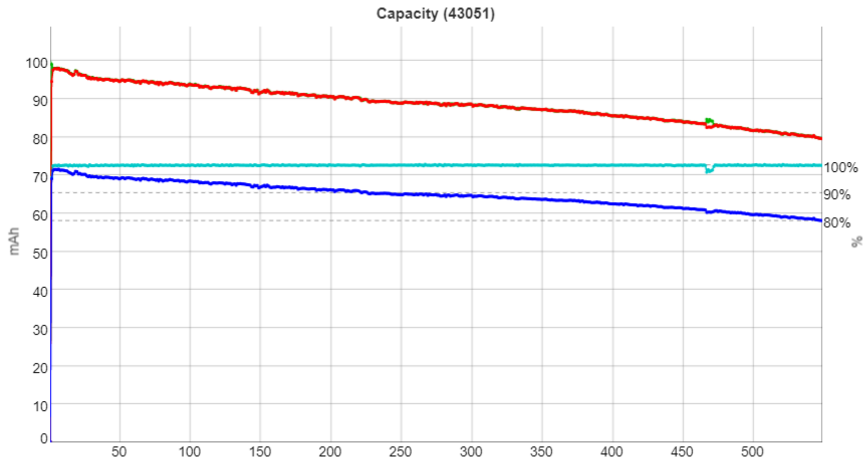
Data Preparation

For IP reasons, some of the data was changed, distorted and reduced. All defective cells that are actually failed experiments were cleaned, data contained valid experiments only. For most of the visualizations (except one) 'currently running' experiments were filtered, data contained 'Finished' experiments only. I used SQL to query specific sub databases, products, experiments, groups and projects. There are many complicated measured parameters in each experiment, for simplicity I focused on only few of them.

Data Visualizations In Use

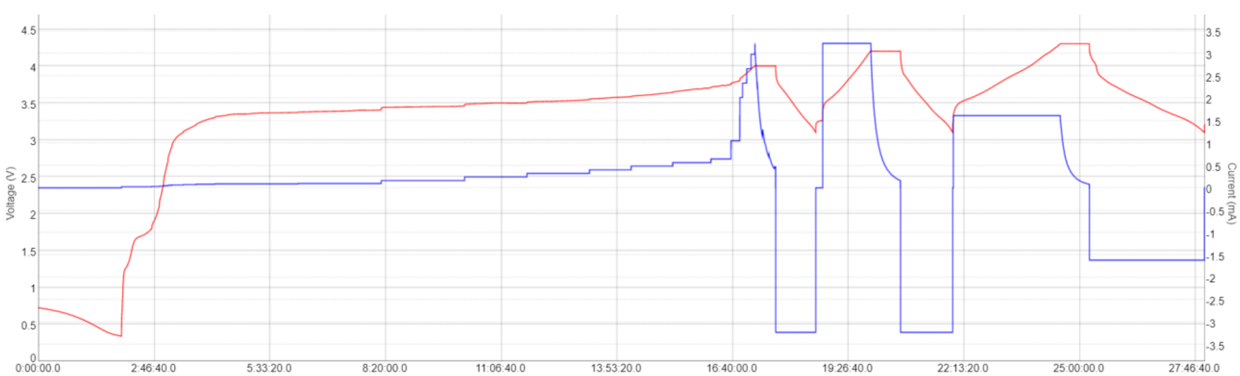
Many of the visualizations in use are for a single cell test. The following figure represents the most common and simple visualization of cell capacity vs the number of charge-discharge cycles. This visualization allows us to inspect the cell's performance during its lifetime.

The left Y axis represents the absolute charge (green) and discharge (red) capacities. The right Y axis represents the normalized discharge capacities (blue) and charge-discharge efficiency (light blue). The X axis represents the cycle number.

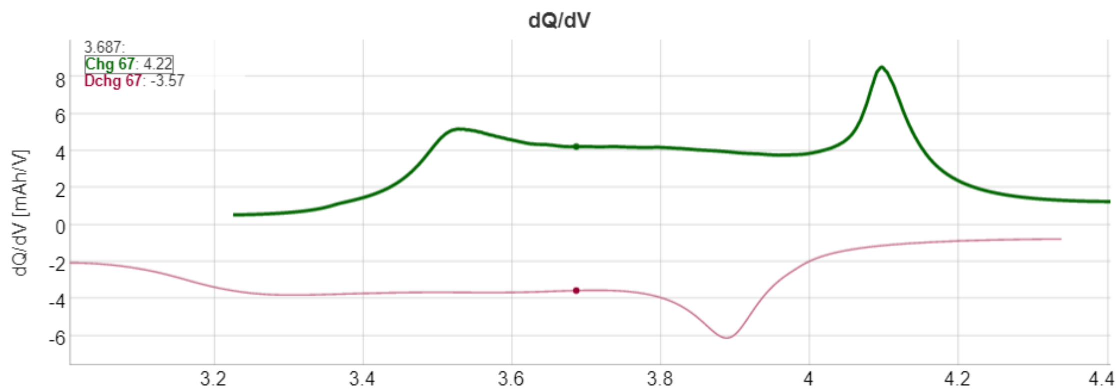


Another popular single cell visualization is the current-voltage curve, where we can zoom-in to any chosen time or cycle number in the test. This visualization allows us to inspect running procedures and profiles, study how different applied conditions affect the battery.

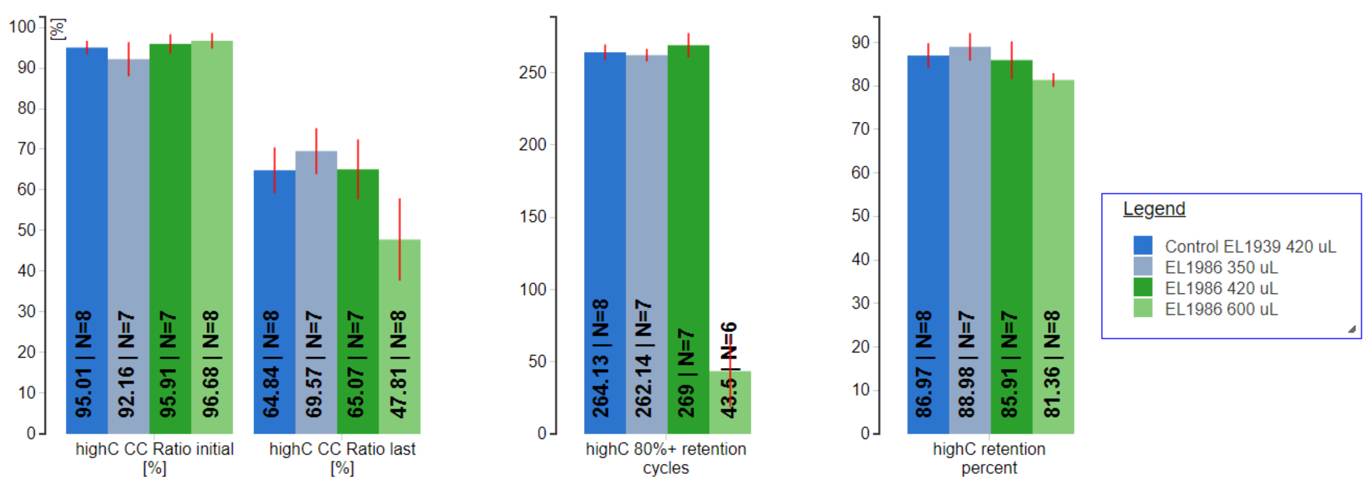
The left Y axis represents the voltage (red) and the right Y axis represents current (blue). The X axis represents time in seconds.



Some visualizations focus on single cycle analysis. The following figure represents the derivative of the voltage and capacity for a particular charge-discharge cycle (cycle 67 in this example). This visualization allows us to inspect chemical reactions and phase transitions in specific voltage values.
The Y axis represents charge (green) and discharge (red) derivative capacities. The X axis represents the voltage.

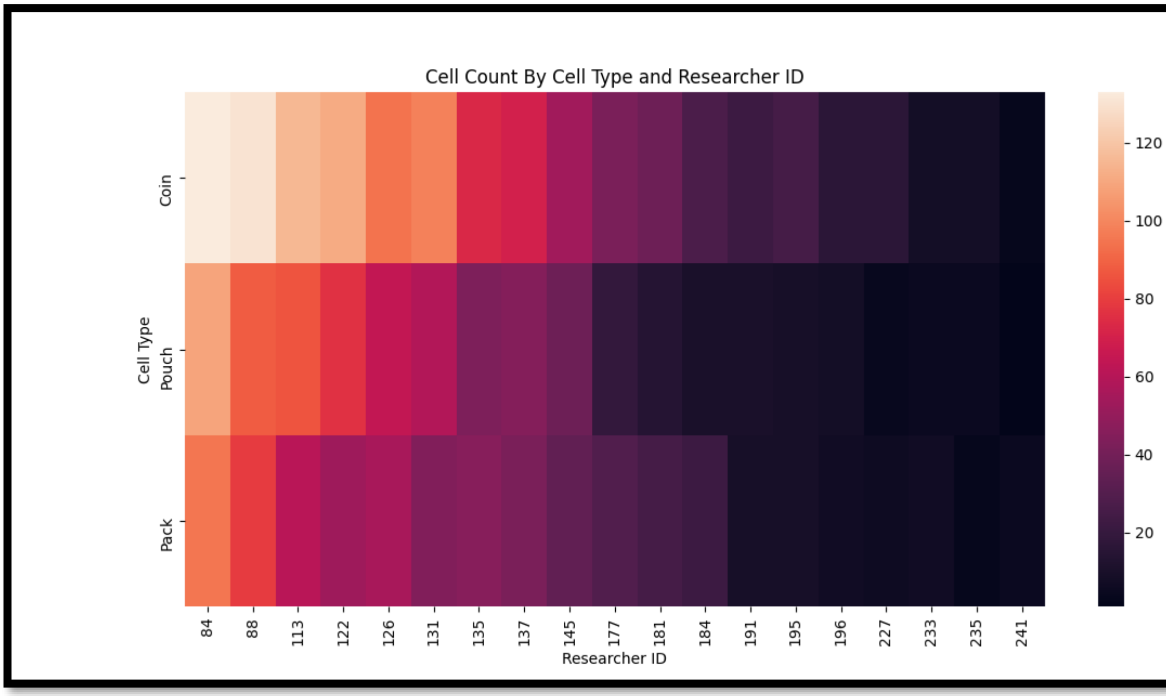


For aggregated data, the tool in use is very simple and the most common chart type is bar chart with error bars. The following figure represents few performance parameters of an electrolyte experiment that tested 7-8 identical cells per experiment point.



Recently, StoreDot started an integration with **Grafana** - a multi-platform open source analytics and interactive visualization web application.

Visualization I – HeatMap



What?

<u>Two categorical key attributes:</u> ‘Cell Type’ and ‘Researcher ID’

<u>One quantitative value attribute:</u> ‘Cell Count’

Each battery cell that is being tested is a part on an experiment that has an owner – a researcher. Every researcher has an id number.
Every battery configuration tested in StoreDot can be produced in three different cell types. The larger the battery the closer it is to representing the final product and it requires more production resources. Most often, configurations in the initial development stage will be tested first in coin cells and will slowly progress to larger cell types.
Data of a specific research group was used to create this visualization.

Coin

Pouch

Pack

How?

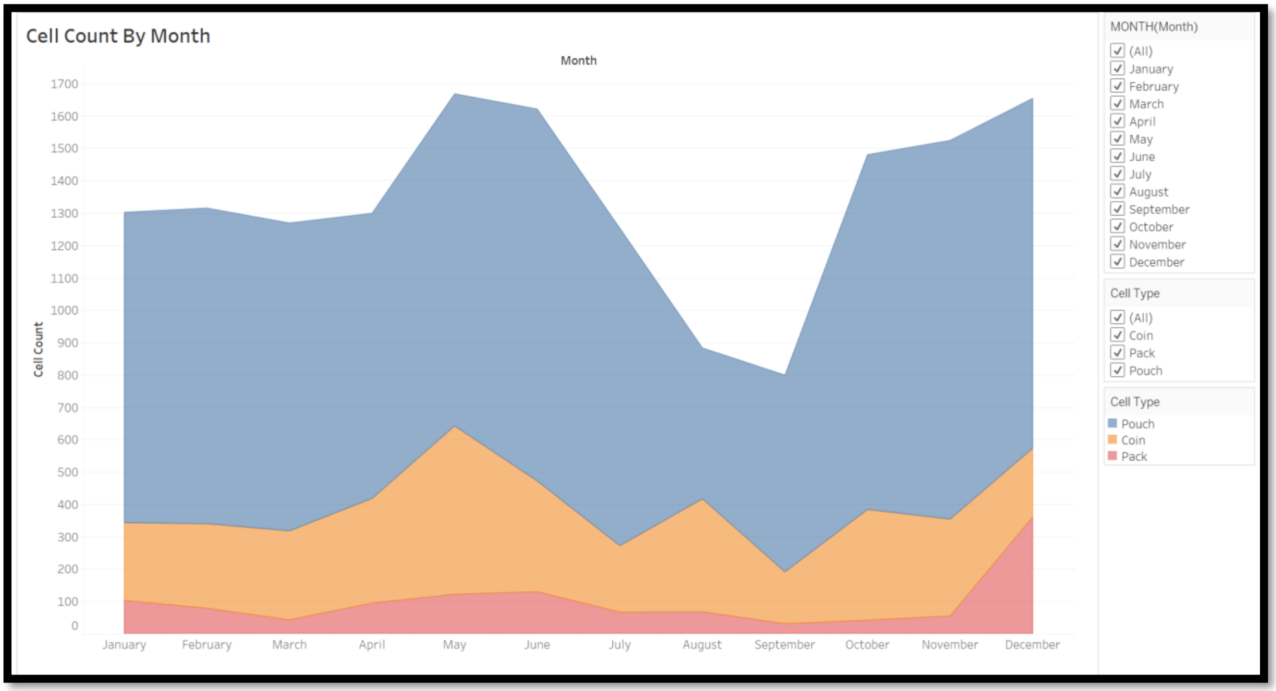
Python - seaborn.heatmap() on pivot table

Why?

Heatmap was chosen because It can represent 3D data in a non-numeric way, so it’s easier to identify patterns and outliers.

From this visualization, we can learn that in general the more experienced the researcher is (lower ID number) the more cells they own. Also, researchers perform more experiments on smaller cell types.

Visualization II - Stacked Area Chart



What?

<u>Quantitative value attribute:</u> ‘Cell count’

<u>Ordered key attribute:</u> ‘Month’

<u>Categorical key attribute:</u> ‘Cell type’

Data of the entire R&D from 2021 was used to create this visualization.

How?

Tableau – area chart with ‘Months’ as dimension, ‘Cell Count’ as measurement and marked by ‘Cell Type’.

Stacks were deliberately ordered from the smaller to the grater area.

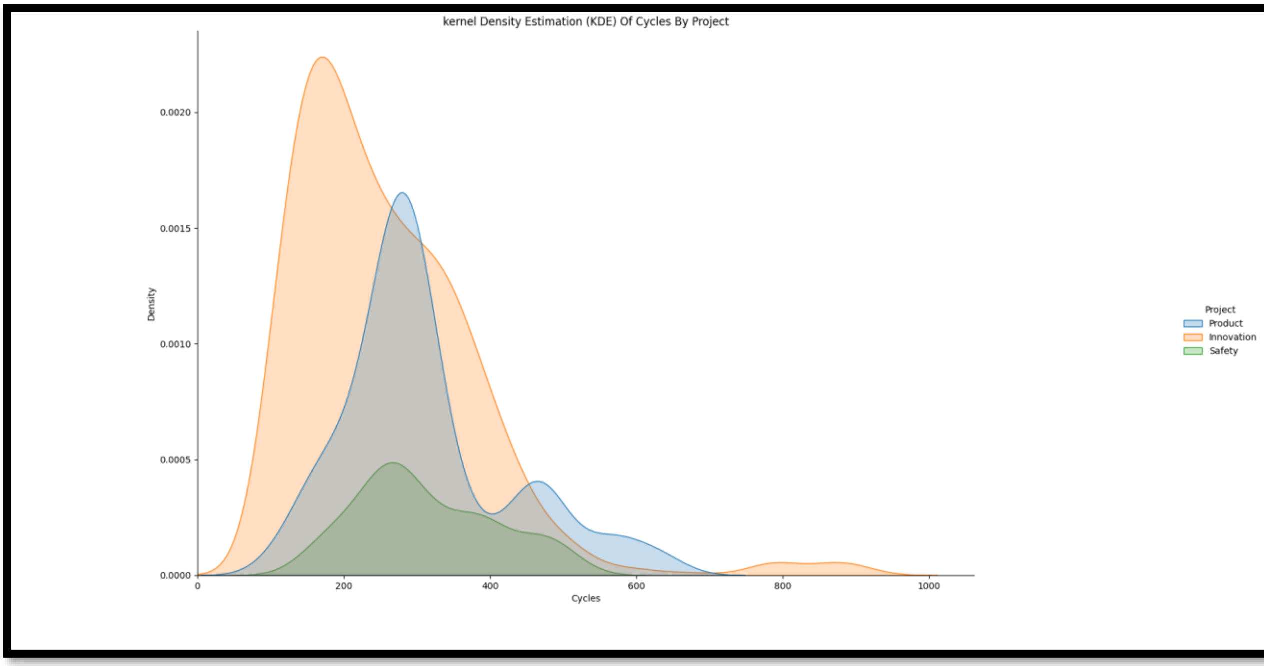
Why?

An interactive stacked area chart was chosen because it allows us to study the evolution of the whole and the relative proportions of each group.

From this visualization, we can learn about trends in R&D cell production over the last year.

How each cell type contributed to the total count and how it compares to others.

Visualization III - Density Curve



What?

<u>Quantitative value attribute:</u> ‘Cycle’

<u>Categorical key attribute:</u> ‘Project

There are three main projects currently running in the company, each of them has different goals in their experiments: **(1) Product** – current baseline performance testing. **(2) Innovation** - future improvements to define new and improved baseline. **(3) Safety -** Current baseline safety testing. Each project focuses on a different stage in the R&D process and therefore exhibits different maturity level when it comes to cycle count distribution.

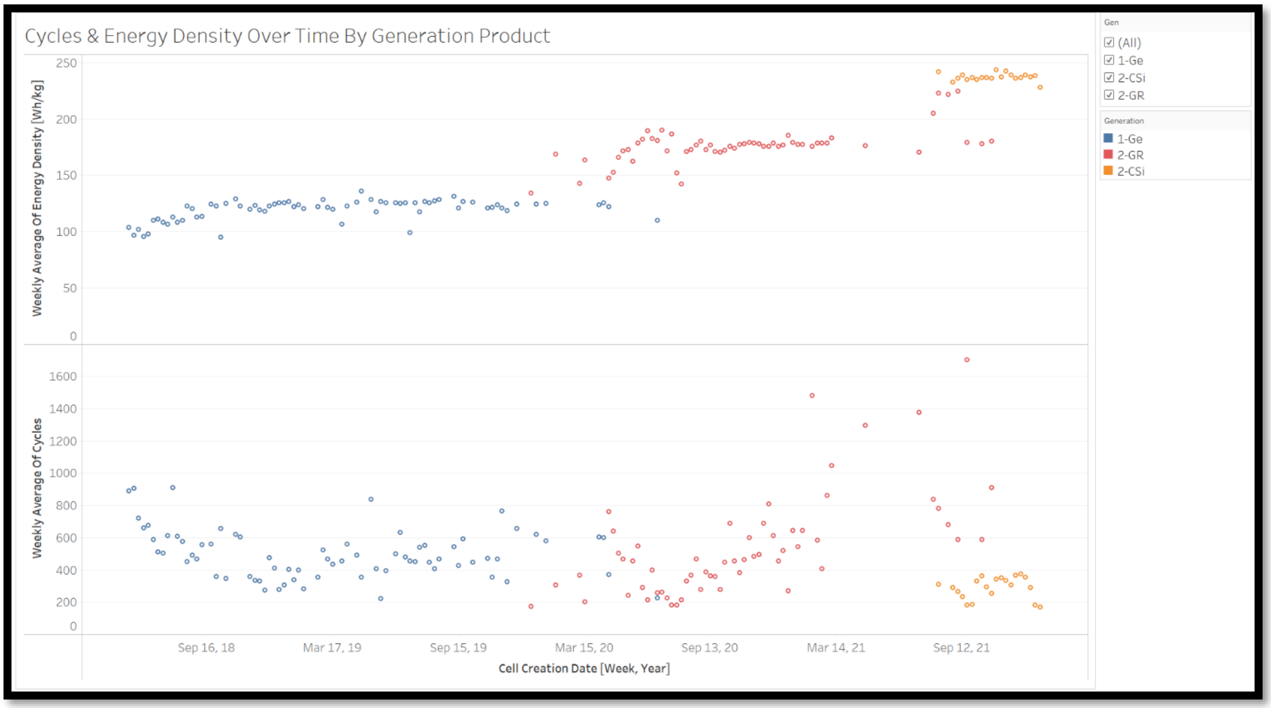
How?

Python - seaborn.displot(kind=‘kde’)

Why?

A density curve chart was chosen to compare distributions instead of amounts to get a more generalized view of each project’s performance.

Visualization IV - Dot Chart



What?

<u>Quantitative value attribute:</u> ‘Cycle’

<u>Quantitative value attribute:</u> ‘Energy Density’

<u>Ordered key attribute:</u> ‘Week’

<u>Categorical attribute:</u> ‘Product generation’

Energy density is another key parameter of cell performance. It usually has an inverse correlation to the number of cycles.
Over the years, StoreDot developed 3 generations of products with different anode materials.

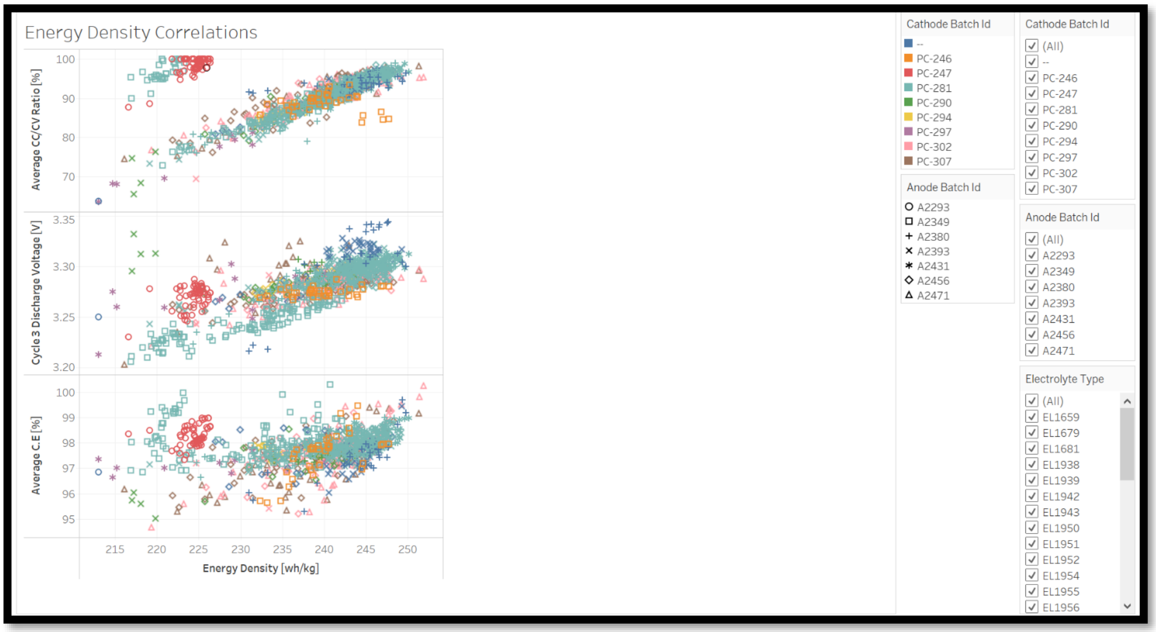
How?

Tableau – dot chart with ‘Week’, ‘AVG Energy Density’, ‘AVG Cycles’ as measurements and marked by ‘Generation’.

Why?

Dot chart was chosen because it’s an easy way to display trends over time.
From this visualization we can learn about StoreDot’s performance progress over time and the relationship between cycles and energy density for each generation.

Visualization V - Scatter Plot Matrix



What?

<u>Target Quantitative value attribute:</u> ‘Energy Density

<u>Quantitative value attributes:</u> ‘‘AVG CC/CV Ratio’, ‘Cycle 3 Discharge Voltage’ & ‘Average C.E’

<u>Categorical attributes:</u> ‘Anode Batch ID’, ‘Cathode Batch ID’ and ‘Electrolyte Type’.

The chosen quantitative attributes are measured during a QC phase before the actual performance testing. The categorical attributes define the cell configuration for the experiment. Each data point represents a single cell.

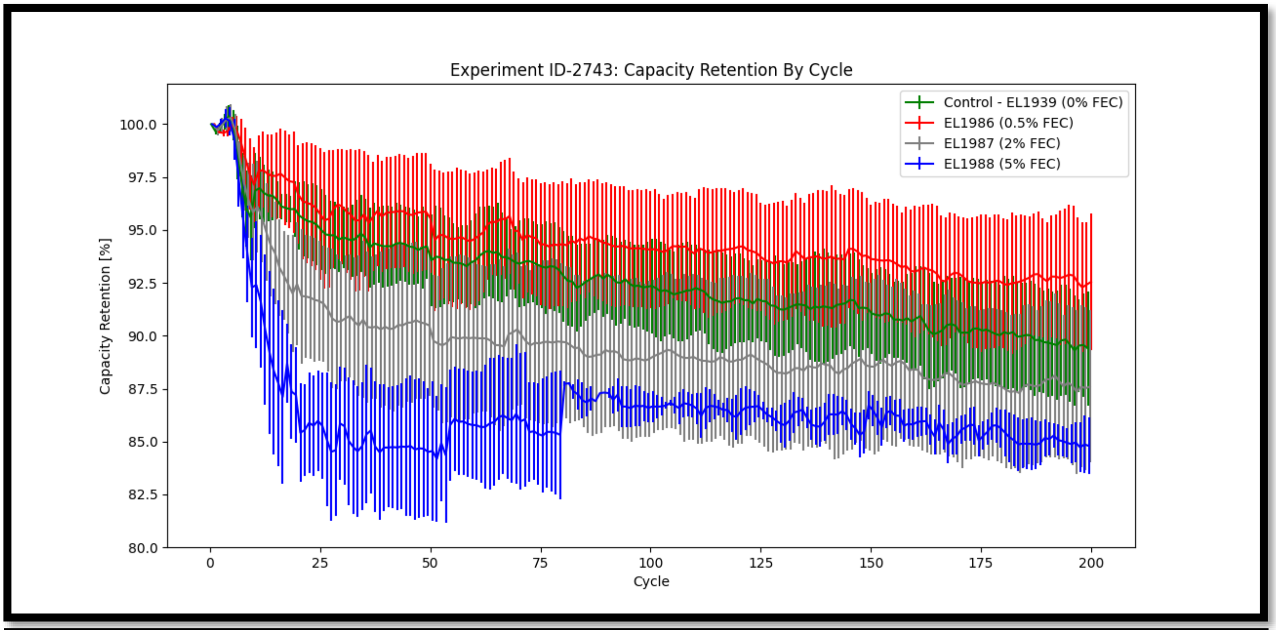
How?

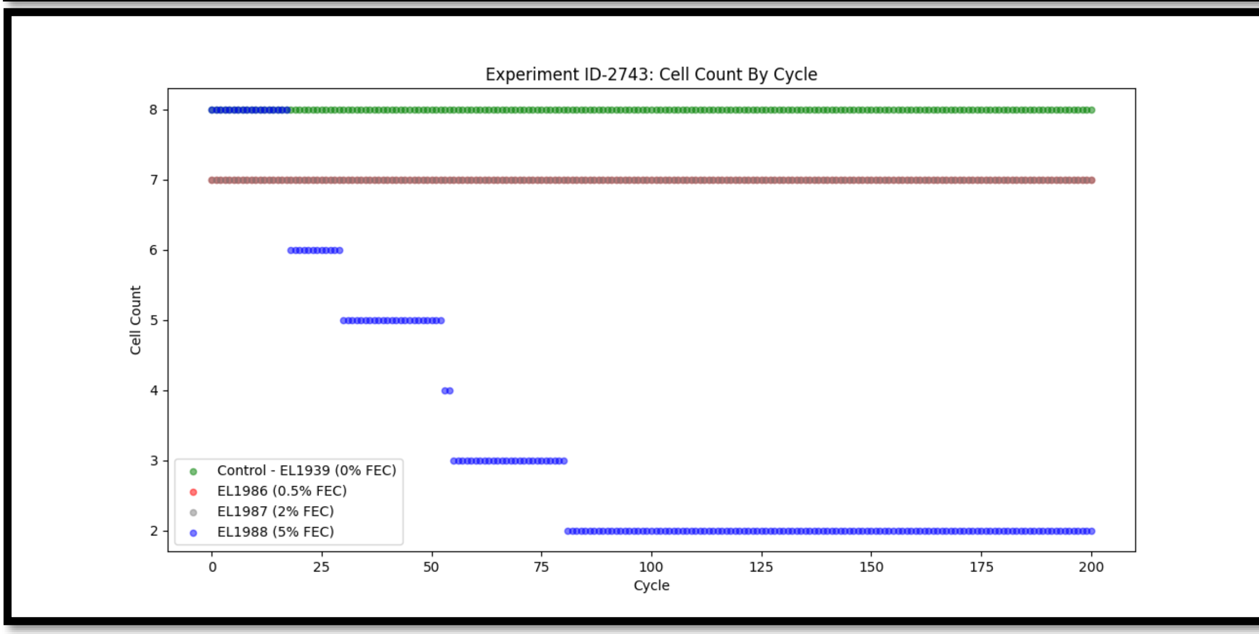
Tableau – scatter plot with ‘Energy Density’, ‘AVG CC/CV Ratio’, Cycle 3 Discharge Voltage’ and ‘Average C.E’ as measurements, marked by ‘Anode Batch ID’, ‘Cathode Batch ID’ and ‘Electrolyte Type’.

Why?

An interactive scatter plot matrix was chosen because it graphically clusters data by attributes. It allows us to detect correlations and associate anomalies to certain attributes.
From this visualization we can learn about correlative QC parameters to the cell’s measured energy density. We can Identify outliers and perform investigation regarding their source. For example – cathode batch id PC-247 (marked in red) shows a significant deviation from the expected values.

Visualization VI - Error Bar





What?

<u>Quantitative value attribute:</u> ‘Capacity Retention’

<u>Quantitative value attribute:</u> ‘Cycle’

<u>Categorical key attribute:</u> ‘Electrolyte Type’

<u>Quantitative value attribute:</u> ‘Cell count’

Data of a “currently running” electrolyte experiment was used to create this visualization. This specific experiment is testing different concentrations of an electrolyte additive called FEC. Each experiment point consists 7-8 cells in identical configuration. The tested performance parameters are the average and standard deviation of the normalized capacity per cycle number. For each cell, the experiment is terminated when the capacity retention is lower than 80%. ‘Cell Count’ refers to the number of cells that reached a certain number of cycles before terminating. For example – in EL1988 (5% FEC) group of cells, only 2 cells out of 8 that ran above 75 cycles. That is why the error bars are dramatically changing whenever a cell is terminated.

How?

Python – Matplotlib.pyplot.errorbar()

Why?

An error bar chart was chosen because in this kind of experiment we are interested in both average and standard deviation of the performance measurement for each experiment group. From this visualization we can get an early indication (while the experiment is still running) on capacity retention – EL1986 (0.5% FEC) currently seem to perform best, since it’s capacity slope is the most moderate. However, its standard deviation is larger than the control group of cells, which harms the significance of this observation.