

# OMICs data analysis in Qlucore: Plots in focus

Yana Stackpole, PhD  
Training, Application Support

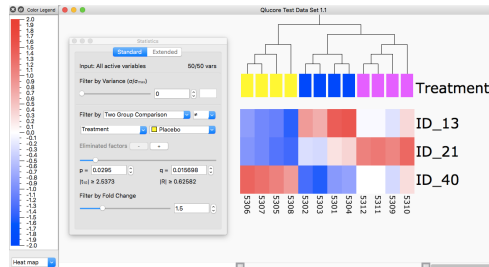
## Visualize and Explore

- QC (outliers, mislabeled samples)
- Make observations - identify structures, patterns
- Generate new hypotheses
- Browse the genome



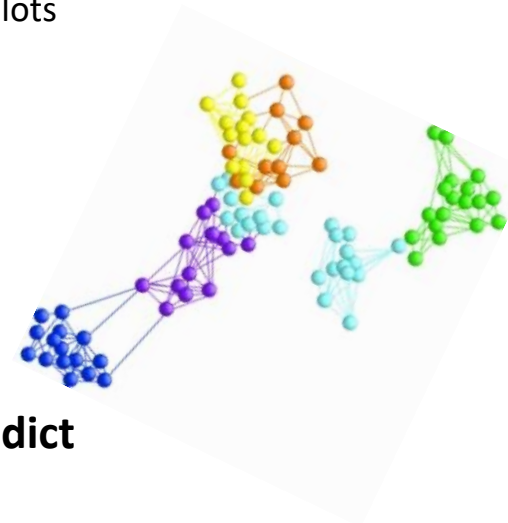
## Analysis

- t-test, ANOVA, Regressions, R scripts. Open API to R, Batch exec
- Variant calling
- Easy generation and export of reports, status, and plots
- Save your session, share



## Biological Insight

- GSEA using MSigDB, reactome, or custom gene sets
- GO Browser
- GO enrichment
- NDEx



## AI - Classify and Predict

- Build classifiers
- kNN, SVM, RT
- Predict sample class, outcome, etc.

# Enjoy fantastic computing speed on a laptop to boost your discovery and scientific creativity

Benchmark examples (static). Compared to R	Times faster
ANOVA (22k var. + 130 samples)	2800
t-test (two-groups, selected from 22k var. + 130 samples)	1000
Kruskal-Wallis (22k var. + 130 samples)	900
Mann-Whitney U-test (two groups, 30k var. + 5k samples)	480
ANOVA (30k var. + 5000 samples)	180
PCA calculations (30k var. + 150 samples)	77
UMAP (22k var. + 130 samples)	13



The speed enables a more flexible workflow – generating better results faster.

Details at: <https://glucore.com/calculation-benchmarks>

**"This tool might literally save you years of your life"**

Prof. Ulrich Steidl at Albert Einstein College of Medicine

# Plot options

The screenshot shows the Qlucore Omics Explorer 3.9 interface. The main menu includes 'File', 'Window', 'License', and 'Help'. The 'Data' tab is active, and the 'Method' sub-tab is selected. The 'Plot' section is highlighted, showing options for 'PCA', 'Line', 'Heat', 'Genome', 'Scatter', 'Bar', 'Table', 'Special', 'All', and 'Variable'. The 'All' option is currently selected. A dropdown menu is open, listing various plot types such as 'PCA Sample', 'PCA Variable', 'Scatter Sample', 'Scatter Variable', 'Line Sample', 'Line Variable', 'Histogram Sample', 'Histogram Variable', 'Bar Sample', 'Box Sample', 'Box Variable', 'Violin Sample', 'Violin Variable', 'Heat', 'Table', 'Volcano', '2D t-SNE', '2D UMAP', 'Kaplan-Meier', 'Silhouette', 'Venn', 'Talus', 'Scree', 'Pie Chart Sample', 'Pie Chart Variable', 'Fusions Circular', and 'Genome'. The 'PCA Variable' option is selected in the dropdown. In the bottom right corner, there is a handwritten note '(5 %)' with a line pointing to the 'All' option in the plot menu.

**Toolbox**

- Multi.
- Move
- Color
- List
- Cent.
- X Axis
- Info
- Mark
- Flip
- X-Val.
- Y Axis
- Annot.
- Label
- Corr.

**Launch**

**Plot**

- PCA
- Line
- Heat
- Genome
- Scatter
- Bar
- Table
- Special
- All
- Variable

**Mode**

- Sample
- Variable

**Normalization**

- None
- Mean=0

**Samples** | **Variables** | Log | NGS

**Variables**

**Lists**

<input type="checkbox"/>	<input type="checkbox"/>	Search	0 0
<input type="checkbox"/>		<b>Acute Lymphobla...emia_Collapsed</b>	<b>100 100</b>
<input type="checkbox"/>	<input checked="" type="checkbox"/>	anova subtype top 100	100 100
<input type="checkbox"/>	<input checked="" type="checkbox"/>	high variance_top 100	100 100
<input type="checkbox"/>	<input checked="" type="checkbox"/>	T ALL vs others, top 100	100 100
<input type="checkbox"/>	<input checked="" type="checkbox"/>	HALLMARK_P53_PATHWAY	200 193

**Dropdown Menu:**

- PCA Sample
- PCA Variable
- Scatter Sample
- Scatter Variable
- Line Sample
- Line Variable
- Histogram Sample
- Histogram Variable
- Bar Sample
- Box Sample
- Box Variable
- Violin Sample
- Violin Variable
- Heat
- Table
- Volcano
- 2D t-SNE
- 2D UMAP
- Kaplan-Meier
- Silhouette
- Venn
- Talus
- Scree
- Pie Chart Sample
- Pie Chart Variable
- Fusions Circular
- Genome

(5 %)

# Venn Diagram

John Venn introduced it in an 1880 paper entitled:

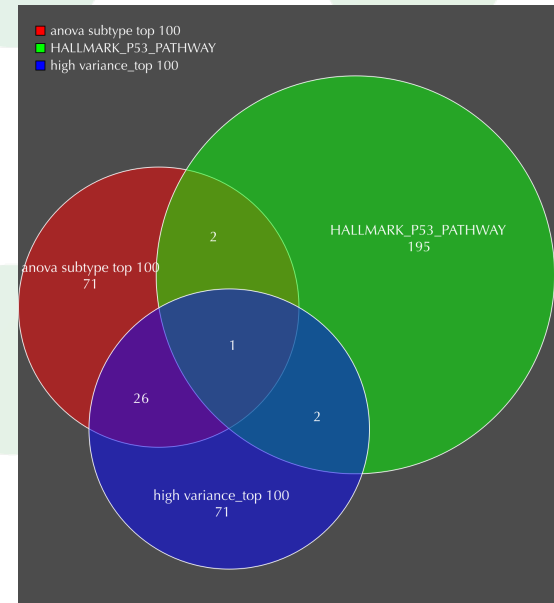
[“On the Diagrammatic and Mechanical Representation of Propositions and Reasonings”](#)

**Math:** Venn diagrams are used for **sets, intersections, and unions**

**Statistics and probability:** Venn diagrams are used to express the **relationships between two or more data sets** in an easily understood way – efficient connections between things

**Simple + Powerful, especially:**

- to highlight similarities and differences, and
- to compare and contrast the characteristics of different sets.



# When?

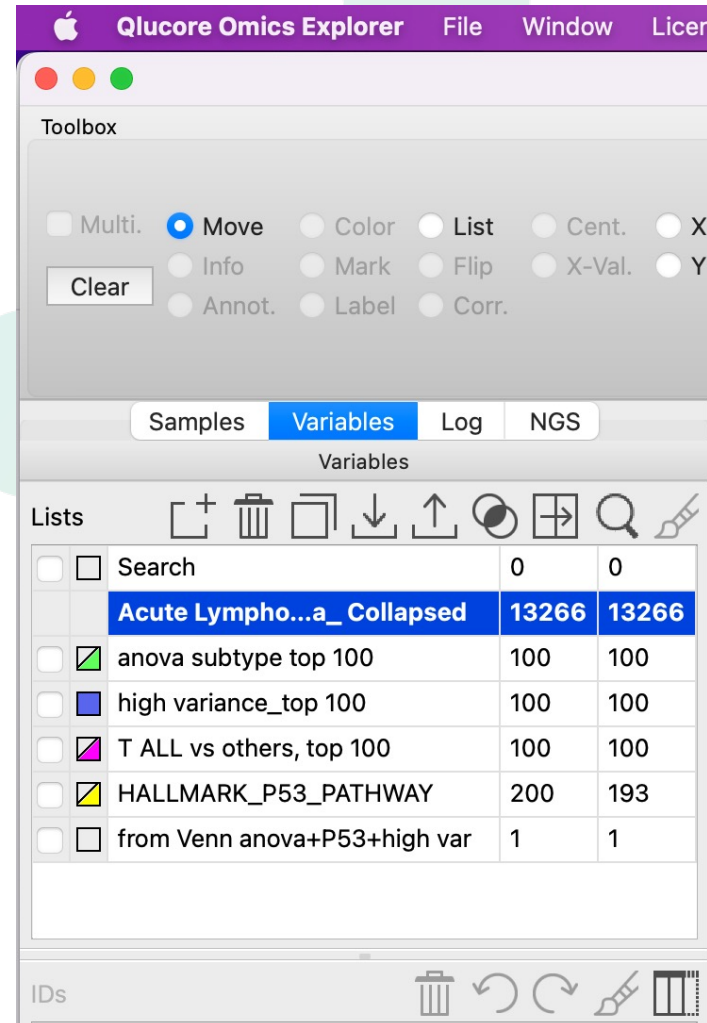
- Venn's initial rationale for it: "Some of X is Y and some of Y is X."  
If *this* is what you're trying to communicate, then Venn.
- "These sets have more (or less) in common than you might think."
- "We're talking about a *very specific* subset."

# Examples of Sets/Lists

## 1. Whole data sets

## 2. Data subsets:

- Statistically significant results
- Variance-based lists
- Pathways, GO lists
- Any lists



The screenshot shows the Qlucore Omics Explorer interface. The 'Variables' tab is active, displaying a list of data subsets. The 'Acute Lympho...a\_Collapsed' subset is highlighted in blue, showing 13266 variables in both columns. Other subsets include 'anova subtype top 100', 'high variance\_top 100', 'T ALL vs others, top 100', 'HALLMARK\_P53\_PATHWAY', and 'from Venn anova+P53+high var'.

Search		0	0
<input checked="" type="checkbox"/>	Acute Lympho...a_Collapsed	13266	13266
<input checked="" type="checkbox"/>	anova subtype top 100	100	100
<input type="checkbox"/>	high variance_top 100	100	100
<input checked="" type="checkbox"/>	T ALL vs others, top 100	100	100
<input checked="" type="checkbox"/>	HALLMARK_P53_PATHWAY	200	193
<input type="checkbox"/>	from Venn anova+P53+high var	1	1

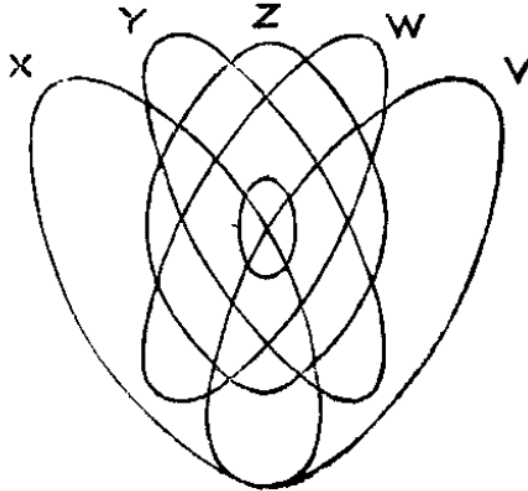
# Venn Diagram

How many sets (terms, gene lists, data sets/subsets) can be used?

See example with 5 sets/terms below (original paper)

x and w, but are no part of z).

It must be admitted that such a diagram is not quite so simple to draw as one might wish it to be; but then we must remember what are the alternatives before any one who wishes to grapple effectively with five terms and all the thirty-two possibilities which they yield. He must either write down or in some

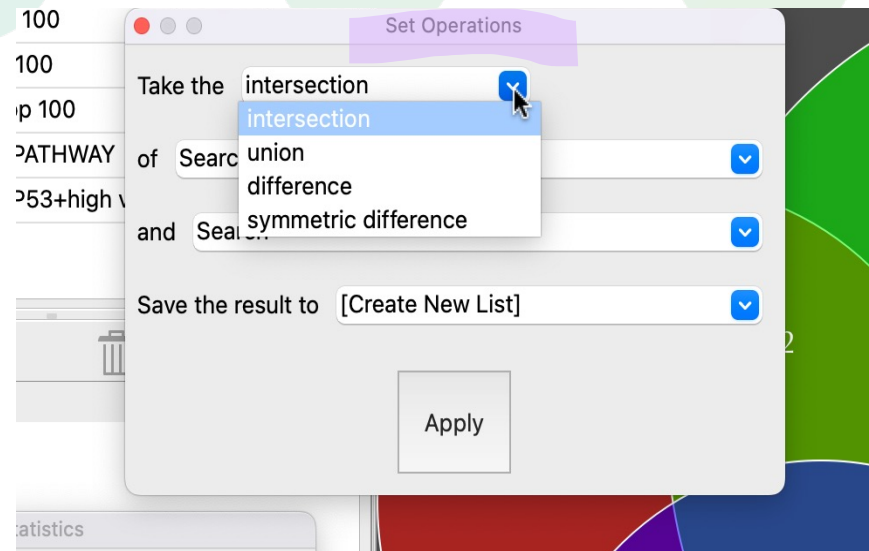




# Handling sets two at a time?

Yes! Not visual, manual, tedious, more operator error prone:

- **Intersection** of A and B
- **Difference** A vs B (unique to A)
- **Symmetric Difference** (unique to A and B merged)
- **Merge sets** without duplicates



## What makes the Qlucore software unique?

- Instant visual feedback
- Easy: to learn – to use – to remember how to use
- Automated access to public data as GEO, TCGA
- Easy import of other public data in a matrix format – like GREIN
- Share analyses, save your sessions to resume later 😊

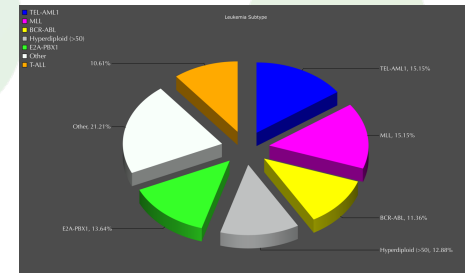
# Pie Chart – part-to-whole relationship in your data

- Displays **relative proportions** of multiple classes of data (% , frequency).
- Size of the circle can be made **proportional** to the total quantity it represents. Hence, can serve as a visual QC check of calculation accuracy.
- Summarizes a **large data set in visual form** – overview of classes.
- Visually simpler than other types of graphs.

**1<sup>st</sup> known Pie** – 1801 Scottish writer on political economy William Playfair (1759-1823), known as the inventor of statistical graphs.

“Making an appeal to the eye when proportion and magnitude are concerned, is the best and readiest method of conveying a distinct idea”

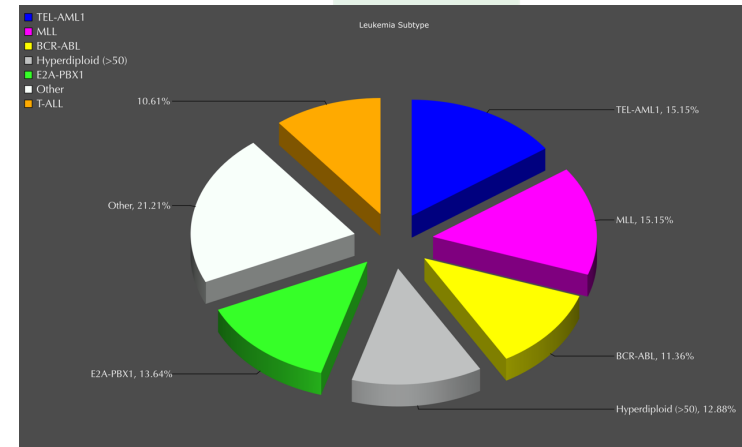
The pie chart and bar chart were described.  
Some believe that he also invented the line plot



# Pie Chart – part-to-whole relationship in your data

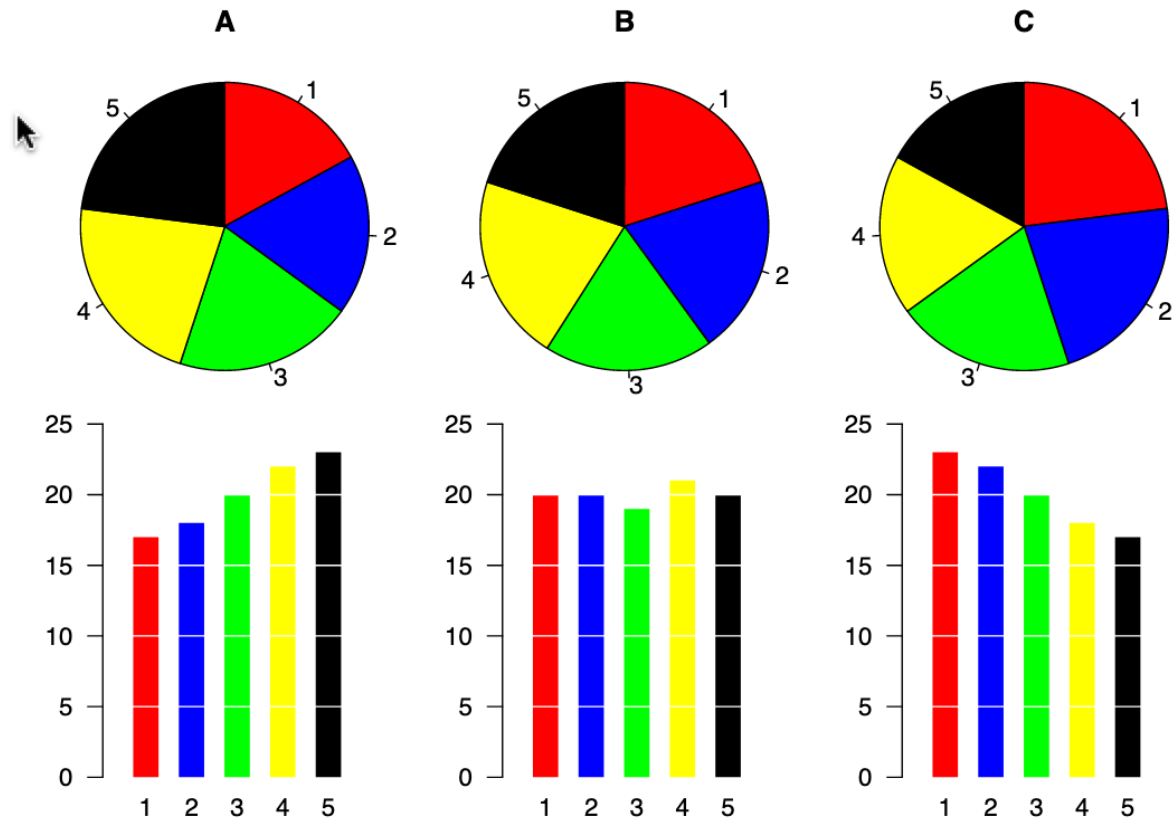
## Keep in mind:

- **Too similar slices** (unless you need to show that, but really no need in a plot). Simple bar plot is better then.
- **Exploded plot may distort proportions**, but helps with many slices.
- **Too many slices**.
- Use of **color**.



# Pie vs Histogram

## “evil pie”



# DEMO

## Next steps

### Experience your data in depth:

1. Join free fully supported trial -- download and install the trial version ([Mac or Win version - links](#)), and then update the license key in **Qlucore – License – Activation Key** with the key from us
2. Book a session with us – get help setting up your analysis.  
<https://calendly.com/yana-stackpole/30min>
3. Link to join our Discord channel with instructional videos and 30 min webinars:  
<https://discord.gg/zu8h8FJx>

Reach out any time [Yana.Stackpole@Qlucore.com](mailto:Yana.Stackpole@Qlucore.com)

# System requirements

## Base module

### *FOR WINDOWS*

- Windows 7, Windows 8 or Windows 10
- 512 MB of RAM memory
- A graphical card with support of at least Open GL 2.1
- 5 GB of free hard disk space
- The program takes full advantage of processors with multiple cores and computers with multiple processors.

### *FOR MAC*

- Max OS X 10.15 or 10.14
- 512 MB of RAM memory
- A graphical card with support of at least Open GL 2.1
- 5 GB of free hard disk space