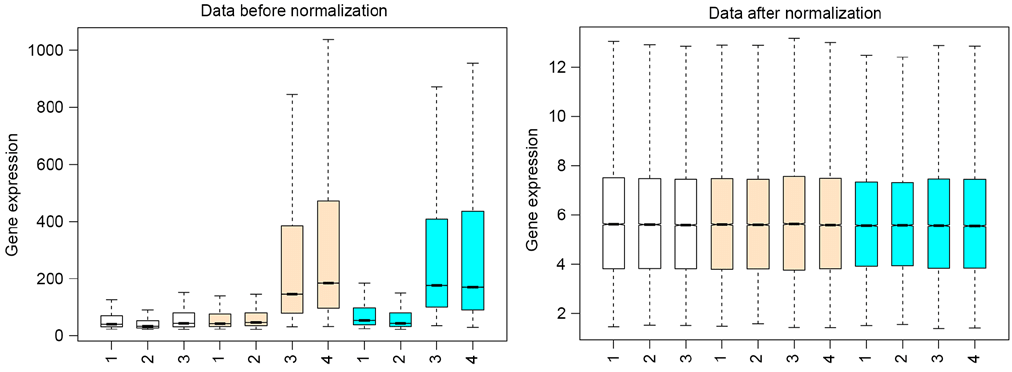
**Zadanie1** Zbadano poziom ekspresji genów u 11 osób przydzielonych do jednej z trzech kategorii (oznaczonych na rysunkach różnymi kolorami). Oś Y reprezentuje poziom ekspresji genów dla kolejnych próbek. Po lewej stronie znajdują się dane przed normalizacją, po prawej – po normalizacji. Wymień i krótko opisz prawdopodobne kroki jakie przeprowadzono.



Wyjaśnij czy sensownym biologicznie jest aby próbki po normalizacji miały zbliżone średnie?

# *NCBI (GEO DataSets) udostępnia dane z eksperymentów mikromacierzowych a także pozwala na ich analizę online. W tym kroku poddamy analizie wybrany taki eksperyment.*

**Zadanie2** Przypadek do analizy z *NCBI GEO DataSets*

* Wejdź w link poniżej:

<https://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS810>

* Czego dotyczyło badanie?
* Gdzie badano ekspresje genów? (jaki materiał/tkanka)
* Ile było wszystkich próbek i na ile grup były one podzielone***(Experiment design and value distribution oraz*** [***Sample Subsets***](https://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS810#details) ***)***
* Według jakich kryteriów podzielono próbki na grupy? Opisz wykorzystane parametry. ***(Experiment design and value distribution oraz*** [***Sample Subsets***](https://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS810#details) ***)***
* Ile jest genów różnicujących? (***Find genes, Find genes that are up/down***

***for this condition(s)***)

* Jak wygląda przebieg ekspresji dla genów *SPARC, VSNL1* oraz *COL5A2* w kolejnych grupach? *(****[Expression Profiles](https://www.ncbi.nlm.nih.gov/geoprofiles?term=GDS810%5bACCN%5d" \t "_blank" \o "View Expression Profiles)*** lub ***Find genes)*** Za co odpowiadają te geny? Czy obserwujesz jakieś tendencje zmiany poziomu ich ekspresji w kolejnych grupach? Poszukaj w źródłach zewnętrznych informacji na temat ich związku z chorobą Alzheimera.
* Czym są *housekeeping genes?* Jaką pełnią rolę w eksperymencie mikromacierzowym? Wybierz trzy przykładowe geny tej kategorii i sprawdź ich ekspresje w kolejnych próbkach.

**Zadanie3** Wyznaczono poziom ekspresji genów dla 6 kolejnych chwil czasowych. Wyznacz współczynnik korelacji pomiędzy poziomem ekspresji genu A oraz genu. Zinterpretuj uzyskany wynik.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Gen 1 | 0.7 | 0.74 | 0.86 | 0.90 | 1.2 | 1.31 |
| Gen 2 | 1.5 | 1.32 | 1.11 | 1.00 | 0.72 | 0.60 |

Możesz użyć poniższy kod Python (zawarte w pliku zadanie3.py).

*Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie*

***STRING***

*Baza danych STRING ma na celu zbieranie, ocenianie i integrowanie wszelkich publicznie dostępnych źródeł informacji na temat interakcji między białkami, a także uzupełnienie ich o prognozy komputerowe. Jej celem jest stworzenie kompleksowej i obiektywnej globalnej sieci, obejmującej zarówno bezpośrednie (fizyczne), jak i pośrednie (funkcjonalne) interakcje białkowe.*

**Zadanie 4** Wejdź na [STRING: functional protein association networks (string-db.org)](https://string-db.org/) a następnie przeanalizuj zestaw potencjalnych genów markerowych dla prognozy raka piersi. Wybierz *Multiple proteins*, a jako organizm *Homo sapiens*.

EFNA1

EGFR

ERBB2

GATA3

GZMB

MST1

MYB

MYBL2

MYC

PLAT

SOX4

SOX9

SRF

XBP1

- Jakie 3 *procesy biologiczne* mają najmniejszy *FDR (false discovery rate)* w rozważanej grupie genów?

- Która *funkcja molekularna* ma najmniejszy *FDR (false discovery rate)* w rozważanej grupie genów?

- Która *ścieżka KEGG* ma najmniejszy *FDR (false discovery rate)* w rozważanej grupie genów?

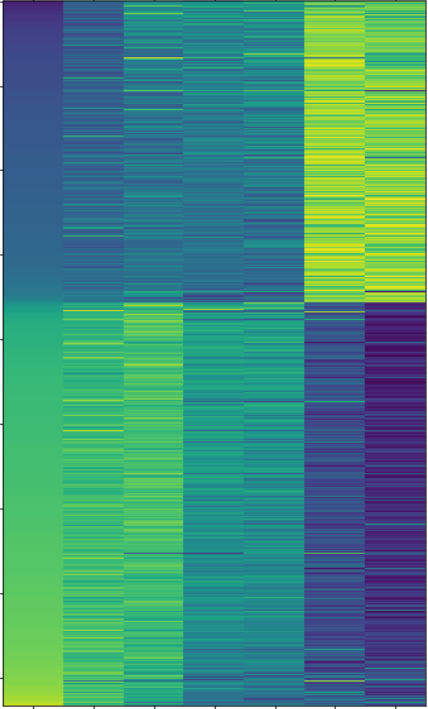
**Zadanie 5** Poniżej zamieszczono dane odnośnie ekspresji genów w komórkach drożdży podczas procesu oddychania – fermentacji alkoholowej. Wyróżniamy dwa główne etapy tego procesu:

- rozkład glukozy do kwasu pirogronowego,

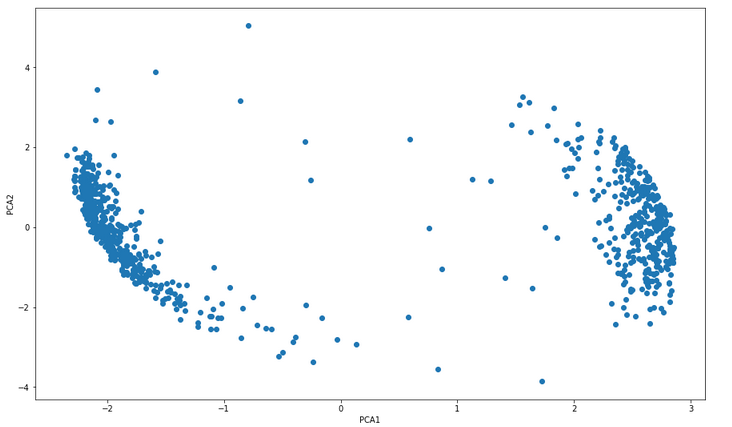
- przemianę kwasu pirogronowego do alkoholu.

Każdy z etapów kontrolowany jest przez 2 klasy genów odpowiedzialnych za te procesy. Dane pochodzą z 7 chwil czasowych (kolejne kolumny). Skomentuj i porównaj poniższe wyniki w kontekście powyższych informacji. W jaki sposób podzieliłbyś rozważane chwile czasowe?

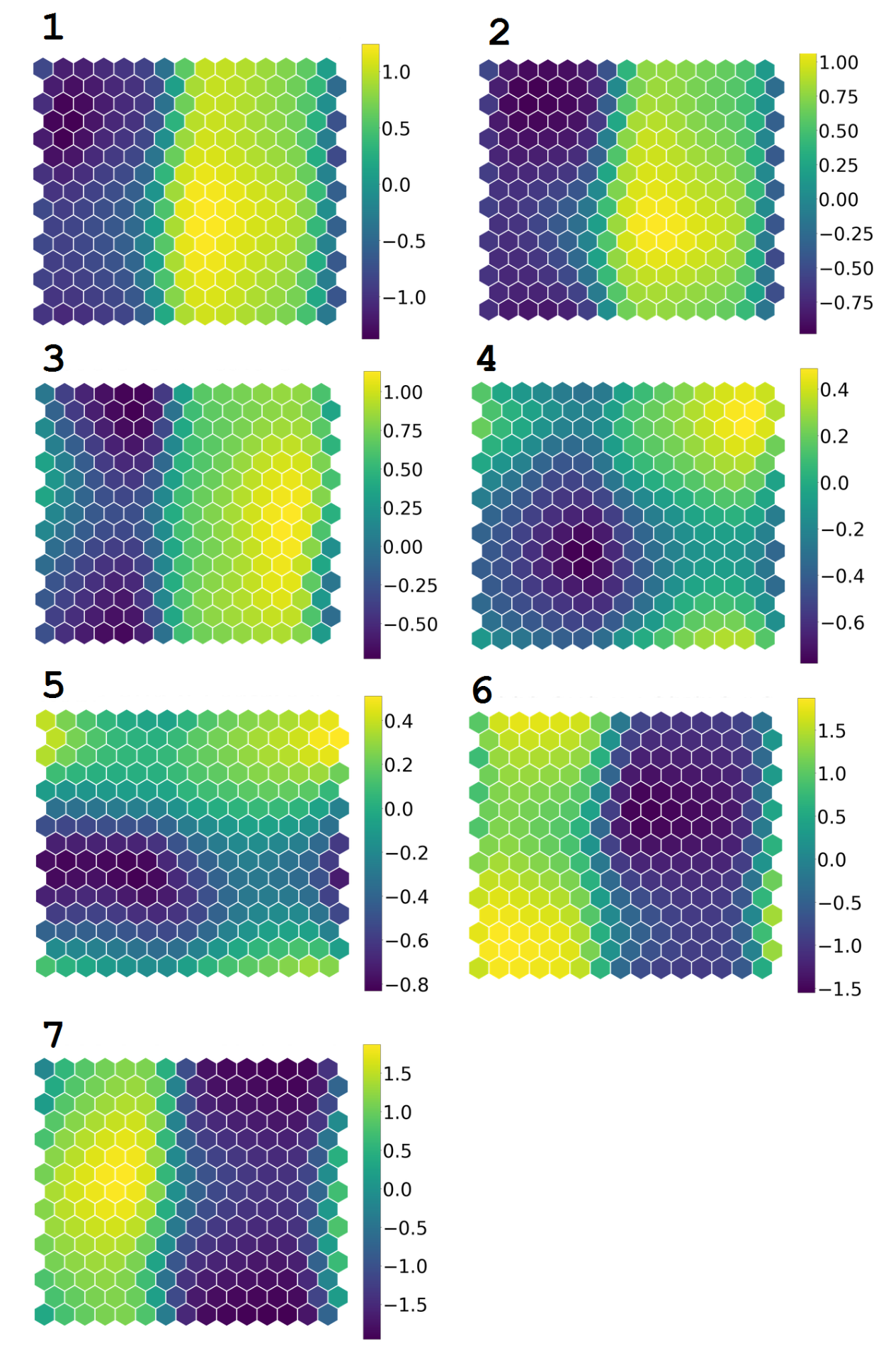
- rozważana mikromacierz (geny zostały posortowane względem pierwszej chwili czasowej):



- po zastosowaniu algorytmu PCA, gdzie jako kolejne obserwacje wybrano wiersze (geny) z powyższej macierzy:



- mapy Kohonena dla kolejnych chwil czasowych. Kolor wskazuje na poziom ekspresji określonej grupy genów:



**Zadanie6** Wejdź na stronę <https://singlecell.broadinstitute.org/single_cell/>. Znajdują się tam dane z eksperymentów scRNA-seq.

- Wyszukaj eksperyment: Single-cell transcriptomics of the spinal cord of a severe SMA mouse

- Czego dotyczył eksperyment? (***Summary***)

- Ile komórek i genów rozważano? (**Summary**)

- Ile typów komórek rozważano (**Explore**)

- Sprawdź ekspresję genów: hemoglobiny (HBB) oraz SPARC wśród rozważanych typów komórek. W której grupie występuje największa ich ekspresja? (**Explore, search genes**)