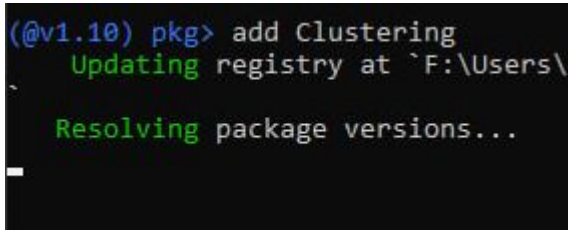
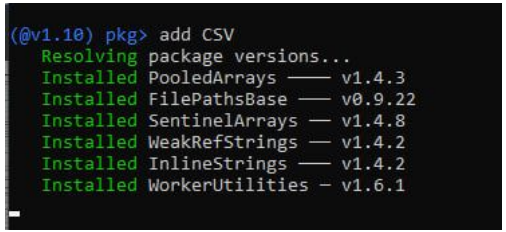
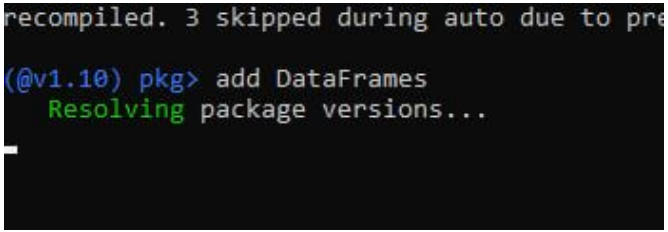
Лабораторная работа №7

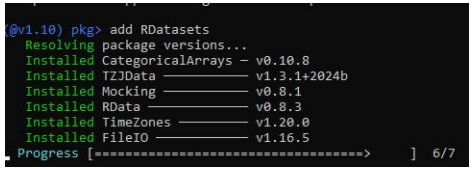
по дисциплине Компьютерный практикум по статистическому анализу данных

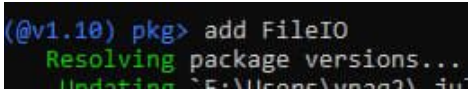
Ход работы

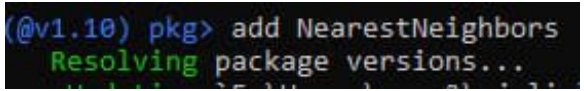

(@v1.10) pkg> add Clustering
Updating registry at `F:\Users\
Resolving package versions...

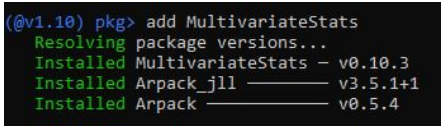

(@v1.10) pkg> add CSV
Resolving package versions...
Installed PooledArrays ─── v1.4.3
Installed FilePathsBase ─── v0.9.22
Installed SentinelArrays ── v1.4.8
Installed WeakRefStrings ── v1.4.2
Installed InlineStrings ─── v1.4.2
Installed WorkerUtilities ─ v1.6.1

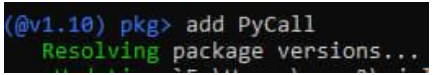

recompiled. 3 skipped during auto due to pre
(@v1.10) pkg> add DataFrames
Resolving package versions...

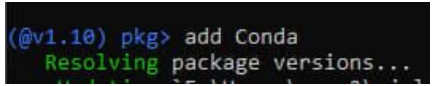

(@v1.10) pkg> add RDatasets
Resolving package versions...
Installed CategoricalArrays ─ v0.10.8
Installed TZJData ────────── v1.3.1+2024b
Installed Mocking ────────── v0.8.1
Installed RData ──────────── v0.8.3
Installed TimeZones ──────── v1.20.0
Installed FileIO ─────────── v1.16.5
Progress [=================================>   ] 6/7

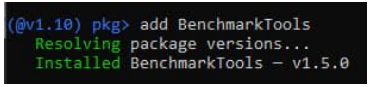

(@v1.10) pkg> add FileIO
Resolving package versions...


(@v1.10) pkg> add NearestNeighbors
Resolving package versions...


(@v1.10) pkg> add MultivariateStats
Resolving package versions...
Installed MultivariateStats ─ v0.10.3
Installed Arpack_jll ──────── v3.5.1+1
Installed Arpack ──────────── v0.5.4


(@v1.10) pkg> add PyCall
Resolving package versions...

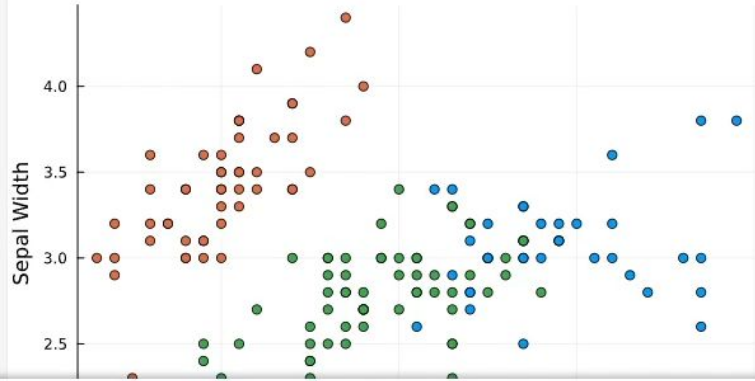

(@v1.10) pkg> add Conda
Resolving package versions...


(@v1.10) pkg> add BenchmarkTools
Resolving package versions...
Installed BenchmarkTools ─ v1.5.0

Ход работы

[31]:
```
# Задание 7.4.1. Кластеризация
using RDatasets, Clustering, Plots, StatsPlots

# Загрузка данных
iris = dataset("datasets", "iris")
X = Matrix(iris[:, 1:4])'  # Транспонирование данных

# Применение метода k-средних
k = 3
result = kmeans(X, k)

# Построение графика кластеров
scatter(X[1, :], X[2, :], group=result.assignments, legend=false, xlabel="Sepal Length", ylabel="Sepal Width")
```

[31]:

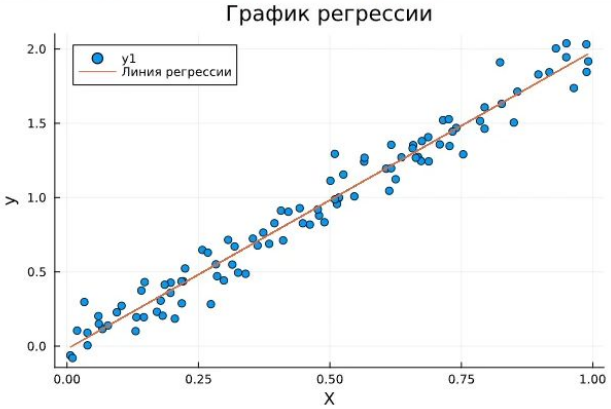


[ ]:

Ход работы

```
y = X * a0 + 0.1 * randn(1000)

# МНК-оценка
X_aug = hcat(ones(1000), X)
beta = X_aug \ y

# Сравнение с llsq из MultivariateStats
model = fit(LinearModel, X, y)
coeff_llsq = coef(model)

# Сравнение с GLM
glm_model = lm(@formula(y ~ x1 + x2 + x3), DataFrame(y=y, x1=X[:,1], x2=X[:,2], x3=X[:,3]))
coeff_glm = coef(glm_model)

# Часть 2
X = rand(100)
y = 2 .* X + 0.1 .* randn(100)

scatter(X, y, xlabel="X", ylabel="y", title="График регрессии")
coeff = [ones(length(X)) X] \ y
plot!(X, coeff[2] .* X .+ coeff[1], label="Линия регрессии")
```

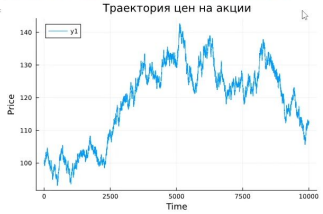
[32]:



Ход работы

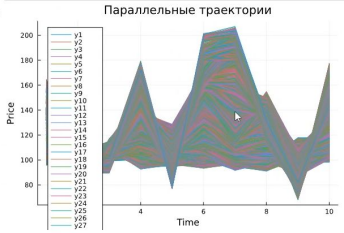
```
    d = exp(r * h - sigma * sqrt(h))
    p_star = (exp(r * h) - d) / (u - d)
    prices = Float64[S]
    for i in 1:n
        push!(prices, prices[end] * (rand() < p_star ? u : d))
    end
    return prices
end

plot()
for i in 1:10
    plot!(createPath(100.0, 0.08, 0.3, 1.0, 10000))
end

# Построение траектории
prices = Float64[S]
for i in 1:n
    push!(prices, prices[end] * (rand() < p_star ? u : d))
end
plot(prices, xlabel="Time", ylabel="Price", title="Траектория цен на акции")
```

[35]:




```
results = pmap(i -> parallelPath(100.0, 0.08, 0.3, 1.0, 10000), 1:10)
plot(hcat(results...)', xlabel="Time", ylabel="Price", title="Параллельные траектории")
```

[36]:

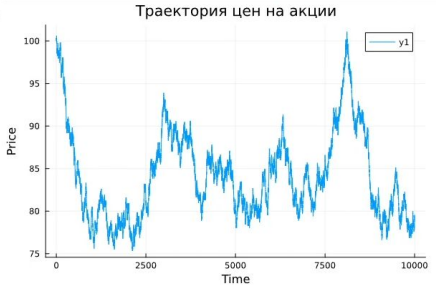


[ ]:

[34]:

```
#Задание 7.4.3. Биномиальные опционы
using Plots

# a)
S = 100.0
T = 1.0
n = 10000
sigma = 0.3
r = 0.08
h = T / n
u = exp(r * h + sigma * sqrt(h))
d = exp(r * h - sigma * sqrt(h))
p_star = (exp(r * h) - d) / (u - d)

# Построение траектории
prices = Float64[S]
for i in 1:n
    push!(prices, prices[end] * (rand() < p_star ? u : d))
end
plot(prices, xlabel="Time", ylabel="Price", title="Траектория цен на акции")
```

[34]:



Вывод

Я освоил специализированные пакетов **Julia** для обработки данных.