

# Package: AFR (via r-universe)

June 21, 2026

**Type** Package

**Title** Toolkit for Regression Analysis of Kazakhstan Banking Sector Data

**Version** 0.3.9

**Copyright** The Agency of the Republic of Kazakhstan for Regulation and Development of Financial Market (ARDFM)

**Maintainer** Sultan Zhaparov <saldau.sultan@gmail.com>

**Description** Tool is created for regression, prediction and forecast analysis of macroeconomic and credit data. The package includes functions from existing R packages adapted for banking sector of Kazakhstan. The purpose of the package is to optimize statistical functions for easier interpretation for bank analysts and non-statisticians. The package also provides helper functions for loading an insurance scoring dataset, a past case competition dataset for insurance risk scoring and fair pricing.

**License** GPL-2

**Depends** R (>= 3.5.0)

**Imports** car, forecast, zoo, olsrr, stats, lmtest, graphics, nlme, ggplot2, tseries, gridExtra, utils, rlang, xts, nortest, goftest, cli, arrow

**Suggests** knitr, rmarkdown, roxygen2

**Encoding** UTF-8

**VignetteBuilder** knitr

**LazyData** true

**Config/roxygen2/version** 8.0.0

**NeedsCompilation** no

**Author** Timur Abilkassymov [aut], Shyngys Shuneyev [aut], Alua Makhmetova [aut], Sultan Zhaparov [aut, cre]

**Config/pak/sysreqs** cmake make libicu-dev libuv1-dev libssl-dev zlib1g-dev

**Repository** <https://ardfm.r-universe.dev>

**Date/Publication** 2026-06-21 18:55:56 UTC

**RemoteUrl** <https://github.com/cran/AFR>

**RemoteRef** HEAD

**RemoteSha** 0294d79e9b47ef35a24727ba794100562455aaee

## Contents

bg	2
bp	3
check_betas	4
checkdata	5
corsel	5
dec_plot	6
difflog	6
finratKZ	7
gq	8
HP	9
load_ogpo_data_dictionary	9
load_ogpo_dataset	10
load_ogpo_test_final	10
load_ogpo_test_with_target	11
load_ogpo_train	12
macroKZ	12
ols_test_normality	15
opt_size	15
pct1	16
pct4	16
pt_multi	17
pt_one	17
reg_plot	18
reg_test	19
regsel_f	19
vif_reg	20
<b>Index</b>	<b>21</b>

---

bg

*Breusch-Godfrey test [BG test]*

---

## Description

BG test is used to test for autocorrelation in the errors of a regression model

**Usage**

```

bg(
  model,
  order = 1,
  order.by = NULL,
  type = c("Chisq", "F"),
  data = list(),
  fill = 0
)

```

**Arguments**

model	is a (generalized)linear regression model
order	integer. maximal order of serial correlation to be tested.
order.by	Either a vector z or a formula with a single explanatory variable like ~ z
type	the type of test statistic to be returned
data	an optional data frame containing the variables in the model
fill	starting values for the lagged residuals in the auxiliary regression. By default 0 but can also be set to NA.

**References**

Mitchel, D. and Zeileis, A. Published 2021-11-07. lmtest package

**Examples**

```

model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
bg(model)

```

---

bp	<i>Breusch-Pagan test</i>
----	---------------------------

---

**Description**

Breusch-Pagan test is used to test against heteroskedasticity of a time-series

**Usage**

```

bp(model, varformula = NULL, studentize = TRUE, data = list())

```

**Arguments**

model	is a (generalized)linear regression model
varformula	a formula describing only the potential explanatory variables for the variance (no dependent variable needed). By default the same explanatory variables are taken as in the main regression model.
studentize	logical. If set to TRUE Koenker's studentized version of the test statistic will be used.
data	an optional data frame containing the variables in the model

**References**

Torsten, H., Zeileis, A., Farebrother, Richard W., Cummins, C., Millo, G., Mitchell, D., lmtest package Wang, B., 2014, bstats package

**Examples**

```
model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
bp(model)
```

---

check_betas	<i>All possible regression variable coefficients.</i>
-------------	---

---

**Description**

Returns the coefficients for each variable from each model.

**Usage**

```
check_betas(object, ...)
```

**Arguments**

object	An object of class lm.
...	Other arguments.

**Value**

check\_betas returns a data.frame containing:

x	model
---	-------

**References**

Hebbali, Aravind. Published 2020-02-10. olsrr package

**Examples**

```
model <- lm(real_gdp~imp+exp+usdkzt+eurkzt, data = macroKZ)
check_betas(model)
```

---

checkdata	<i>Preliminary data check for errors</i>
-----------	--

---

**Description**

Preliminary check of data frame for missing values, numeric format, outliers.

Missing items: The number of missing values in each column of the dataset. Numeric format: The number of non-numeric variables in each column of the dataset. Outliers: The number of outliers in each column of the dataset.

**Usage**

```
checkdata(x)
```

**Arguments**

x is a data frame

**Examples**

```
data(macroKZ)
checkdata(macroKZ)
```

---

corse1	<i>Multicollinearity test</i>
--------	-------------------------------

---

**Description**

multicollinearity is the occurrence of high interrelations among two or more independent variables in a multiple regression.

**Usage**

```
corse1(x, thrs, num)
```

**Arguments**

x is a numeric vector or matrix  
thrs threshold set to calculate correlation above  
num logical

**Examples**

```
data(macroKZ)
corse1(macroKZ, num=FALSE, thrs=0.65)
```

dec\_plot

*Decomposition plot*

---

**Description**

The function depicts decomposition of regressors as a stacked barplot

**Usage**

```
dec_plot(model, dataset, print_plot = TRUE)
```

**Arguments**

model	An object of class lm.
dataset	A dataset based on which model was built
print_plot	logical

**Author(s)**

The Agency of the Republic of Kazakhstan for Regulation and Development of Financial Market (AFR)

**References**

Hebbali, Aravind. Published 2020-02-10. olssr package

**Examples**

```
data(macroKZ)
model <- lm(real_gdp ~ usdkzt + eurkzt + imp+exp, data = macroKZ)
dec_plot(model, macroKZ)
```

---

difflog*Transforming time-series data to stationary*

---

**Description**

Difference of logarithms is finding the difference between two consecutive logarithm values of a time-series

**Usage**

```
difflog(x, lag = 1, difference = 1)
```

**Arguments**

x	time-series vector
lag	lagged period
difference	difference between x items

**Examples**

```
data (macroKZ)
new<-pct1(macroKZ)
```

---

finratKZ	<i>finratKZ dataset</i>
----------	-------------------------

---

**Description**

finratKZ dataset

**Usage**

```
finratKZ
```

**Format**

Dataset of 400 corporate borrowers, i.e. 200 standard (IFRS stage 1) and 200 default ones, characterized by 29 financial ratios.

**Default** Dummy variable where 0 - standard(IFRS stage 1) borrower, 1 - default borrower

**Rev\_gr** Revenue growth rate

**EBITDA\_gr** EBITDA growth rate

**Cap\_gr** Capital growth rate

**CR** Current ratio

**QR** Quick ratio

**Cash\_ratio** Cash ratio

**WC\_cycle** Working capital cycle

**DTA** Debt-to-assets

**DTE** Debt-to-equity

**LR** Leverage ratio (Total assets/Total equity)

**EBITDA\_debt** EBITDA-to-debt

**IC** Interest coverage (Income statement)

**CTI** Cash-to-income

**IC\_CF** Interest coverage (Cash flow statement)

**DCR** Debt coverage ratio (Cash flow from operations/Total debt)

**CFR** Cash flow to revenue  
**CRA** Cash return on assets (Cash flow from operations/Total assets)  
**CRE** Cash return on equity (Cash flow from operations/Total equity)  
**ROA** Return on assets  
**ROE** Return on equity  
**NPM** Net profit margin  
**GPM** Gross profit margin  
**OPM** Operating profit margin  
**RecT** Receivables turnover  
**InvT** Inventory turnover  
**PayT** Payables turnover  
**TA** Total assets turnover  
**FA** Fixed assets turnover  
**WC** Working capital turnover

## References

The Agency of the Republic of Kazakhstan for Regulation and Development of Financial Market

---

gq

*Godfrey-Quandt test*

---

## Description

Godfrey-Quandt test is used to test against heteroskedasticity of a time-series

## Usage

```

gq(
  model,
  point = 0.5,
  fraction = 0,
  alternative = c("greater", "two.sided", "less"),
  order.by = NULL,
  data = list()
)
  
```

## Arguments

model	is a (generalized)linear regression model
point	numerical. If point is smaller than 1 it is interpreted as percentages of data
fraction	numerical. The number of central observations to be omitted.
alternative	a character string specifying the alternative hypothesis.
order.by	Either a vector z or a formula with a single explanatory variable like ~ z
data	an optional data frame containing the variables in the model.

**References**

Torsten, H., Zeileis, A., Farebrother, Richard W., Cummins, C., Millo, G., Mitchell, D., Imtest package Wang, B., 2014, bstats package

**Examples**

```
model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
gq(model)
```

---

HP	<i>Hodrick-Prescott filter for time series data</i>
----	---

---

**Description**

Hodrick-Prescott filter is a data smoothing technique that removes trending in time series data frame

**Usage**

```
HP(x, freq = NULL, type = c("lambda", "frequency"), drift = FALSE)
```

**Arguments**

x	time-series vector
freq	integer
type	character, indicating the filter type
drift	logical

**Examples**

```
data(macroKZ)
HP(macroKZ[,2])
```

---

load_ogpo_data_dictionary	<i>Load OGPO data dictionary</i>
---------------------------	----------------------------------

---

**Description**

Loads the OGPO data dictionary included in the package.

**Usage**

```
load_ogpo_data_dictionary()
```

**Value**

A data frame with column names and descriptions.

**Examples**

```
data_dictionary <- load_ogpo_data_dictionary()
head(data_dictionary)
```

---

load_ogpo_dataset	<i>Load all OGPO insurance scoring datasets</i>
-------------------	---

---

**Description**

Downloads and loads all OGPO datasets and the data dictionary.

**Usage**

```
load_ogpo_dataset()
```

**Value**

A list with 'train', 'test\_with\_target', 'test\_final' and 'data\_dictionary'.

**Examples**

```
## Not run:
ogpo_data <- load_ogpo_dataset()
names(ogpo_data)

## End(Not run)
```

---

load_ogpo_test_final	<i>Load OGPO insurance scoring final test dataset</i>
----------------------	---

---

**Description**

Downloads the OGPO final scoring dataset from GitHub Releases if it is not already available in the local cache.

**Usage**

```
load_ogpo_test_final()
```

**Value**

A data frame intended for final scoring and generation of predictions.

**Examples**

```
## Not run:  
ogpo_test_final <- load_ogpo_test_final()  
dim(ogpo_test_final)  
  
## End(Not run)
```

---

```
load_ogpo_test_with_target  
Load OGPO insurance scoring test dataset with target
```

---

**Description**

Downloads the OGPO validation/test dataset with known target values from GitHub Releases if it is not already available in the local cache.

**Usage**

```
load_ogpo_test_with_target()
```

**Value**

A data frame for model validation, quality assessment and business-effect calculation.

**Examples**

```
## Not run:  
ogpo_test_with_target <- load_ogpo_test_with_target()  
dim(ogpo_test_with_target)  
  
## End(Not run)
```

---

load_ogpo_train	<i>Load OGPO insurance scoring training dataset</i>
-----------------	---

---

### Description

Downloads the OGPO training dataset from GitHub Releases if it is not already available in the local cache, then reads it as a data frame.

### Usage

```
load_ogpo_train()
```

### Value

A data frame with policy, driver, vehicle, premium, cancellation, claim and target information.

### Examples

```
## Not run:
ogpo_train <- load_ogpo_train()
dim(ogpo_train)

## End(Not run)
```

---

macroKZ	<i>macroKZ dataset</i>
---------	------------------------

---

### Description

macroKZ dataset

### Usage

```
macroKZ
```

### Format

A time series data frame of 61 quarterly observations of 50 macroeconomic and 10 financial parameters for 2010-2025 period.

**real\_gdp** Real GDP

**GDD\_Agr\_R** Real gross value added Agriculture

**GDD\_Min\_R** Real gross value added Mining

**GDD\_Min\_R** Real gross value added Mining

**GDD\_Man\_R** Real gross value added Manufacture

**GDD\_Elc\_R** Real gross value added Electricity  
**GDD\_Con\_R** Real gross value added Construction  
**GDD\_Trd\_R** Real gross value added Trade  
**GDD\_Trn\_R** Real gross value added Transportation  
**GDD\_Inf\_R** Real gross value added Information  
**GDD\_R** Real gross value added  
**GDP\_DEF** GDP deflator  
**Rincpop\_q** Real population average monthly income  
**Rexppop\_q** Real population average monthly expenses  
**Rwage\_q** Real population average monthly wage  
**imp** Import  
**exp** Export  
**cpi** Inflation  
**realest\_resed\_prim** Real price for estate in primary market  
**realest\_resed\_sec** Real price for estate in secondary market  
**realest\_comm** Real price for commercial estate  
**index\_stock\_weighted** Change in stock value for traded companies  
**ntrade\_Agr** Change in stock value for non-traded companies Agriculture  
**ntrade\_Min** Change in stock value for non-traded companies Mining  
**ntrade\_Man** Change in stock value for non-traded companies Manufacture  
**ntrade\_Elc** Change in stock value for non-traded companies Electricity  
**ntrade\_Con** Change in stock value for non-traded companies Construction  
**ntrade\_Trd** Change in stock value for non-traded companies Trade  
**ntrade\_Trn** Change in stock value for non-traded companies Transportation  
**ntrade\_Inf** Change in stock value for non-traded companies Information  
**fed\_fund\_rate** Federal Funds Rate  
**govsec\_rate\_kzt\_3m** Return on government securities in KZT, 3 m  
**govsec\_rate\_kzt\_1y** Return on government securities in KZT, 1 year  
**govsec\_rate\_kzt\_7y** Return on government securities in KZT, 7 years  
**govsec\_rate\_kzt\_10y** Return on government securities in KZT, 10 years  
**tonia\_rate** TONIA  
**rate\_kzt\_mort\_0y\_1y** Weighted average mortgage lending rate for new loans, less than a year  
**rate\_kzt\_mort\_1y\_1y** Weighted average mortgage lending rate for new loans, more than a year  
**rate\_kzt\_corp\_0y\_1y** Weighted average mortgage lending rate for new loans to non-financial organizations in KZT, less than a year  
**rate\_usd\_corp\_0y\_1y** Weighted average mortgage lending rate for new loans to non-financial organizations in CKB, less than a year

**rate\_kzt\_corp\_1y\_iy** Weighted average mortgage lending rate for new loans to non-financial organizations in KZT, more than a year

**rate\_usd\_corp\_1y\_iy** Weighted average mortgage lending rate for new loans to non-financial organizations in CKB, more than a year

**rate\_kzt\_indv\_0y\_1y** Weighted average mortgage lending rate for consumer loans in KZT, less than a year

**rate\_kzt\_indv\_1y\_iy** Weighted average mortgage lending rate for consumer loans in KZT, less than a year

**usdkzt** USD KZT exchange rate

**eurkzt** EUR KZT exchange rate

**rurkzt** RUB KZT exchange rate

**poil** Price for Brent

**realest\_resed\_prim\_rus** Real price for estate in primary market in Russia

**realest\_resed\_sec\_rus** Real price for estate in secondary market in Russia

**cred\_portfolio** credit portfolio

**coef\_liq\_k4** k4 prudential coefficient

**coef\_k1** k1 prudential coefficient

**coef\_k3** k3 prudential coefficient

**provisions** provisions

**percent\_margin** percent margin

**com\_inc** commissioner income

**com\_exp** commissioner expenses

**oper\_inc** operational income

**oth\_inc** other income

**DR** default rate

## Source

Bureau of National statistics, Agency for Strategic planning and reforms of the Republic of Kazakhstan

## References

The Agency of the Republic of Kazakhstan for Regulation and Development of Financial Market

---

ols\_test\_normality      *Test for normality Test for detecting violation of normality assumption.*

---

**Description**

Test for normality Test for detecting violation of normality assumption.

**Usage**

```
ols_test_normality(model, ...)
```

**Arguments**

model                  an object of class lm.  
 ...                    Other arguments.

**Value**

ols\_test\_normality is a list containing the following components:

kolmogorv	kolmogorov smirnov statistic
shapiro	shapiro wilk statistic
cramer	cramer von mises statistic
anderson	anderson darling statistic

**Examples**

```
data(macroKZ)
model <- lm(real_gdp ~ imp + exp + usdkzt + poil, data = macroKZ)
ols_test_normality(model)
```

---

opt\_size                  *Necessary size of the time-series dataset*

---

**Description**

Estimates number of models generated from given number of regressors X

**Usage**

```
opt_size(model)
```

**Arguments**

model                  is a linear regression model a class lm.

**Examples**

```
data(macroKZ)
model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
opt_size(model)
```

---

pct1

*Transforming time-series data to stationary*

---

**Description**

Percent change is a change between two consecutive terms,

**Usage**

```
pct1(x)
```

**Arguments**

x                    time-series vector(s)

**Examples**

```
data (macroKZ)
new<-pct1(macroKZ)
```

---

pct4

*Transforming time-series data to stationary*

---

**Description**

Percent change is a change between a term and its lagged value for prior period,

**Usage**

```
pct4(x)
```

**Arguments**

x                    time-series vector(s)

**Examples**

```
data (macroKZ)
new<-pct4(macroKZ)
```

---

pt_multi	<i>Pluto-Tasche method for multi-year probability of default (PD) analysis</i>
----------	--

---

**Description**

Calculates the variation inflation factors of all predictors in regression models

**Usage**

```
pt_multi(pf, num_def, conf_level, num_years)
```

**Arguments**

pf	unconditional portfolio distribution from the worst to the best credit quality
num_def	number of defaults in a given rating class
conf_level	confidence interval of PD estimates
num_years	number of periods used in the PD estimation

**Examples**

```
pf <- c(10,20,30,40)
num_def <- c(1,2,3,4)
conf_level = 0.99
num_years = 3
pt_multi(pf, num_def, conf_level, num_years)
```

---

pt_one	<i>Pluto-Tasche method for one-year probability of default (PD) analysis</i>
--------	--

---

**Description**

Calculates probability of default according to One-period Pluto and Tasche model

**Usage**

```
pt_one(pf, num_def, ci = 0.9)
```

**Arguments**

pf	unconditional portfolio distribution from the worst to the best credit quality
num_def	number of defaults in a given rating class
ci	confidence interval of PD estimates

## References

Surzhko, Denis. Published 2015-05-21. LDPD package. Archived on 2022-06-20.

## Examples

```
pf <- c(10,20,30,40)
num_def <- c(1,2,3,4)
pt_one(pf, num_def, ci= 0.9)
```

---

reg\_plot

*Regression forecast plot*

---

## Description

The function depicts forecast and actual data.

## Usage

```
reg_plot(model, dataset)
```

## Arguments

model	An object of class lm.
dataset	A dataset based on which model was built.

## Author(s)

The Agency of the Republic of Kazakhstan for Regulation and Development of Financial Market (AFR)

## Examples

```
data(macroKZ)
model <- lm(real_gdp ~ usdkzt + eurkzt + imp + exp, data = macroKZ)
reg_plot(model, macroKZ)
```

---

reg_test	<i>Test for detecting violation of Gauss-Markov assumptions.</i>
----------	--

---

**Description**

Test for detecting violation of Gauss-Markov assumptions.

**Usage**

```
reg_test(y)
```

**Arguments**

y                    A numeric vector or an object of class lm.

**Value**

reg\_test returns an object of class "reg\_test". An object of class "reg\_test" is a list containing the following components:

bp	Breusch-Pagan statistic
bg	Breusch-Godfrey statistic
dw	Durbin-Watson statistic
gq	Godfrey-Quandt statistic

**Examples**

```
data(macroKZ)
model <- lm(real_gdp~ imp + exp + poil + eurkzt + usdkzt, macroKZ)
reg_test(model)
```

---

regsel_f	<i>Regressors selection</i>
----------	-----------------------------

---

**Description**

The function allows to choose regressors based on multiple criteria as AIC, RMSE etc

**Usage**

```
regsel_f(
  model,
  pval = 0.3,
  metric = "adjr" & "aic",
  progress = FALSE,
  details = FALSE,
  ...
)
```

**Arguments**

model	is a linear regression model
pval	p value; variables with p value less than pval will enter into the model
metric	statistical metrics used to estimate the best model
progress	Logical; if TRUE, will display variable selection progress.
details	Logical; if TRUE, will print the regression result at each step.
...	other arguments

**References**

Hebbali, Aravind. Published 2020-02-10. olssr package

**Examples**

```
data(macroKZ)
model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
regsel_f(model)
```

---

vif\_reg

*VIF by variable*


---

**Description**

Calculates the variation inflation factors of all predictors in regression models

**Usage**

```
vif_reg(model)
```

**Arguments**

model	is a linear regression model
-------	------------------------------

**References**

Petrie, Adam. Published 2020-02-21. regclass package

**Examples**

```
data(macroKZ)
model <- lm(real_gdp ~ imp + exp + poil + eurkzt + tonia_rate, data = macroKZ)
vif_reg(model)
```

# Index

## \* datasets

finratKZ, [7](#)

macroKZ, [12](#)

bg, [2](#)

bp, [3](#)

check\_betas, [4](#)

checkdata, [5](#)

corsel, [5](#)

dec\_plot, [6](#)

difflog, [6](#)

finratKZ, [7](#)

gq, [8](#)

HP, [9](#)

load\_ogpo\_data\_dictionary, [9](#)

load\_ogpo\_dataset, [10](#)

load\_ogpo\_test\_final, [10](#)

load\_ogpo\_test\_with\_target, [11](#)

load\_ogpo\_train, [12](#)

macroKZ, [12](#)

ols\_test\_normality, [15](#)

opt\_size, [15](#)

pct1, [16](#)

pct4, [16](#)

pt\_multi, [17](#)

pt\_one, [17](#)

reg\_plot, [18](#)

reg\_test, [19](#)

regsel\_f, [19](#)

vif\_reg, [20](#)