

# Audit Commander

## Linear regression on data in Access databases

### Audit task in brief

Test relationships on data in Access databases using linear regression and stepwise regression..

### Data source(s)

- Access

### Typical Audit Uses

Preliminary audit work to identify potential areas of audit focus.

### Description of Output

Output is a report produced by the “R” statistical software system.

### Form Image

The screenshot shows a Windows-style application window titled "frmMDBReg - Regression on Access data". The window contains several input fields and controls:

- Access DB:** C:\Program Files\EZS\AC\mdb\fa.mdb
- Report File:** C:\Program Files\EZS\AC\mdb\fa.rep
- Table:** FA (dropdown menu)
- Dep Column:** Asset Cost (dropdown menu)
- Extraction Criteria:** [asset cost] > 5000
- Regression formula:** ATION +ACQDATE +ACCODE +DISPDATE +REPCOST +ACQUISITIO
- Regression Types:** Radio buttons for Linear and Stepwise. The Stepwise option is selected.
- Buttons:** Exit and Process.

# Audit Commander

## Linear regression on data in Access databases

### How to complete the form and run it

Perform the following steps:

1. Click on the input file and identify the Access database from which the data is to be extracted
2. If the name of the report file to contain the resulting report is not acceptable, type in a new name or select a file using the menu..
3. Select the table which contains the data to be analyzed.
4. Select the column values from the drop-down list of the variables to be used for the dependent variable.
5. Specify the extraction criteria, using the column names and values to be tested. An example of an extraction of all records whose net book value is less than salvage would be entered as “(Cost – AD) < Salvage” (without the quotes). Any column name can be tested. Functions (e.g. min, max) can also be included as part of the test.
6. Specify the regression formula to be used. The format is dependent variable name followed by a tilde (~) followed by one or more independent variables separated by a plus sign..
7. Select the regression using the radio button, i.e. linear regression or stepwise linear regression..
8. Click the “OK” button to perform the regression and wait until the “hourglass” goes away, which signifies that processing has been completed.
9. View the report shown in notepad when the processing completes.
10. It is possible to go back and repeat the process using different parameters, etc.
11. When finished, click the “Cancel” button to return to the main menu.

### Example Output

Output is the “R” Report. (sample is shown below). First is for linear regression and the second is for stepwise regression.

#### **Linear Regression**

Call:

lm(formula = Cost ~ AD)

Residuals:

Min	1Q	Median	3Q	Max
-17092.3	-451.7	133.6	758.8	6206.7

Coefficients:

Estimate	Std. Error	t value	Pr(> t )
----------	------------	---------	----------



Audit Commander  
Linear regression on data in Access databases

```
(Intercept) 1.075e+03 4.165e+01 25.81 <2e-16 ***  
AD 1.872e+00 1.614e-02 115.94 <2e-16 ***  
---  
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Residual standard error: 1297 on 3862 degrees of freedom  
Multiple R-squared: 0.7768, Adjusted R-squared: 0.7768  
F-statistic: 1.344e+04 on 1 and 3862 DF, p-value: < 2.2e-16

```
> proc.time()  
user system elapsed  
1.20 0.01 1.20  
.
```

**Stepwise regression**

Start: AIC=55394.81  
Cost ~ AD

Start: AIC=55394.81  
Cost ~ AD

	Df	Sum of Sq	RSS	AIC
<none>		6.4967e+09	5.5395e+04	
- AD	1	2.2614e+10	2.9111e+10	6.1188e+04

Call:  
lm(formula = Cost ~ AD)

Residuals:

Min	1Q	Median	3Q	Max
-17092.3	-451.7	133.6	758.8	6206.7

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.075e+03	4.165e+01	25.81	<2e-16 ***
AD	1.872e+00	1.614e-02	115.94	<2e-16 ***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1297 on 3862 degrees of freedom  
Multiple R-squared: 0.7768, Adjusted R-squared: 0.7768  
F-statistic: 1.344e+04 on 1 and 3862 DF, p-value: < 2.2e-16

Call:



Audit Commander  
Linear regression on data in Access databases

lm(formula = Cost ~ AD)

Residuals:

Min	1Q	Median	3Q	Max
-17092.3	-451.7	133.6	758.8	6206.7

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.075e+03	4.165e+01	25.81	<2e-16 ***
AD	1.872e+00	1.614e-02	115.94	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1297 on 3862 degrees of freedom

Multiple R-squared: 0.7768, Adjusted R-squared: 0.7768

F-statistic: 1.344e+04 on 1 and 3862 DF, p-value: < 2.2e-16

> proc.time()

user	system	elapsed
1.21	0.06	1.23

Audit Commander  
Linear regression on data in Access databases