CH. 13 Linear Regression

RUNNING ANALYSES USING JAMOVI
INTERPRETTING OUTPUT
USING APA STYLE

PY 204 Statistics for Research

Dr. Valenti

Outline for Ch. 13 – Linear Regression

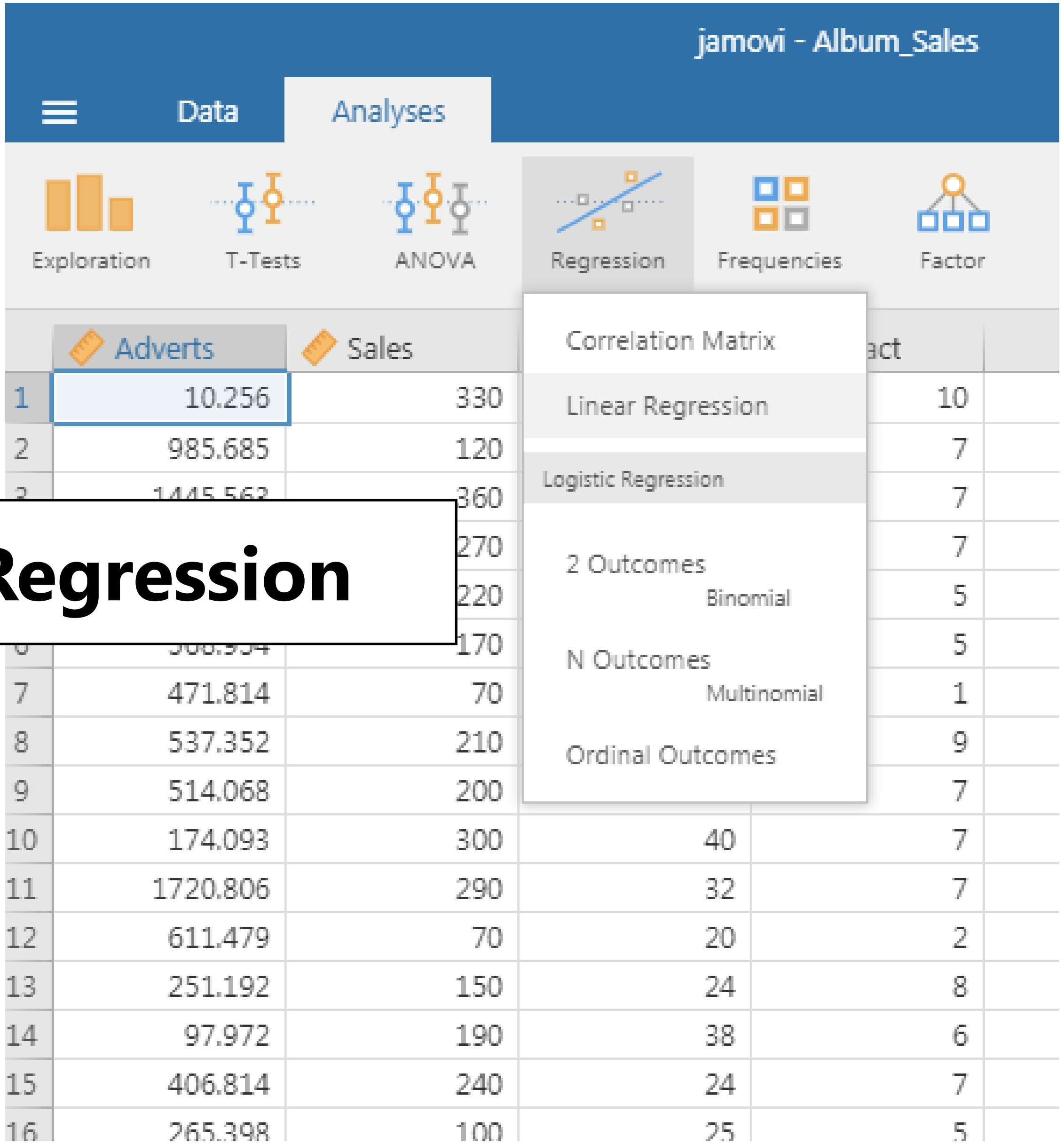
1. **How to run linear regression using JAMOVI**
2. How to interpret JAMOVI output
3. How to write up results using APA style

Regression using JAMOVI: Example Study

- A record company boss was interested in estimating future record sales from advertising money spent.
- Data:
  - 200 different record releases
  - Outcome variable (the var you are *trying to predict*):
    - **# of records sold** in the week after its release
  - Predictor variable (the var you are *predicting from*):
    - The **amount of money** (in £) **spent on advertising** for the record, before its release

What is the predictor variable?
What is the outcome variable?

Regression Using JAMOVI



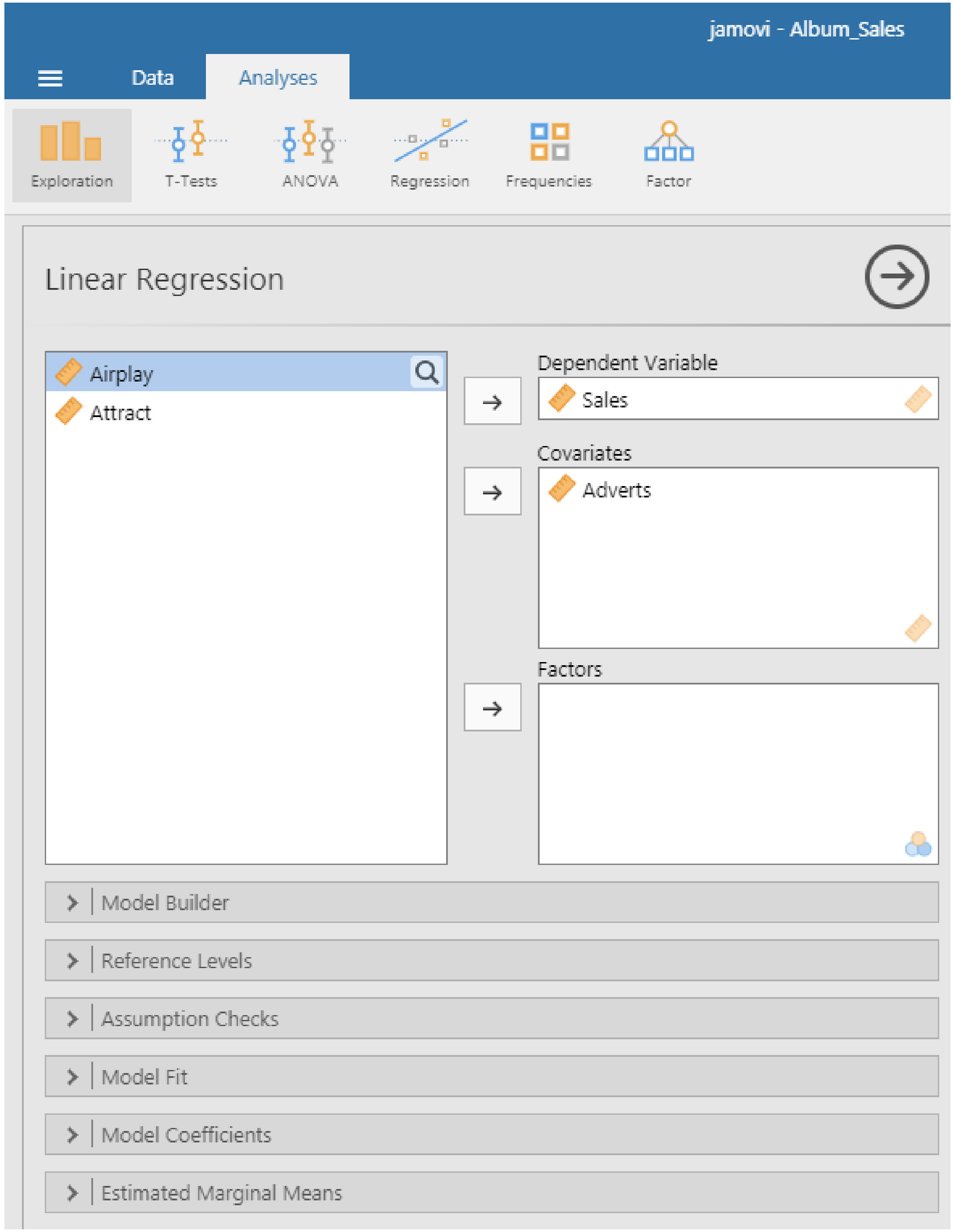
Analyses → Regression → Linear Regression

Regression Using JAMOVI

Move **outcome** variable into "**Dependent Variable**" box.

Move **predictor** variable into "**Covariates**" box

Leave the "Factors" box and everything else as is.



Outline for Ch. 13 – Linear Regression

1. How to run linear regression using JAMOVI
2. **How to interpret JAMOVI output**
3. How to write up results using APA style

null & alternative hypotheses

- In regression, we test a hypothesis about whether individual predictors (Xs) are related to the outcome variable (Y).
  - **null hypothesis** (words): There is *no relationship* between our predictor and outcome.
  - H_0: $b_1 = 0$
  - **alternative hypothesis** (words): There is *some relationship* between our predictor and outcome.
  - H_1: $b_1 \neq 0$

$$\hat{Y} = b_0 + b_1 X_i$$

JAMOVI Output: Model Fit Measures & Model Coefficients

outcome variable (Dependent Variable box)

predictor variable (Covariates box)

effect size

The % of variance in the **outcome** explained by the **predictor.**

33% of the variance in **record sales** is explained by **ad budget**.

Linear Regression

Model Fit Measures

Model	R	R^2
1	0.58	0.33

Model Coefficients - Sales

Predictor	Estimate	SE	t	p
Intercept	134.14	7.54	17.80	< .001
Adverts	0.10	0.01	9.98	< .001

JAMOVI Output

Regression Equation: $Y = b_0 + b_1X$
***aka* eqn for “best fit” line** $Y = 134 + .10X$

intercept (b_0)

t statistic for b_1

p-value for the *t* statistic associated with b_1

Model Coefficients - Sales

Predictor	Estimate	SE	t	p
Intercept	134.14	7.54	~~17.80~~	~~< .001~~
Adverts	0.10	0.01	9.98	< .001

ignore these

regression coefficient, *aka* the slope (b_1)

Interpreting the regression coefficient (slope, b_1)

Regression Equation: $Y = 134 + .10X$
***aka* eqn for "best fit" line**

HOW DO WE INTERPRET THE SLOPE?

Model Coefficients - Sales

Outcome

Predictor→

Predictor	Estimate	SE	t	p
Intercept	134.14	7.54	17.80	< .001
Adverts	0.10	0.01	9.98	< .001

For every 1 unit increase in the predictor (X), the outcome (Y) increases by **.10** units.

Here, for every <u>1 unit</u> increase in ad budget, sales increase by .10 <u>units</u>.

For every <u>£1000</u> increase in ad budget, sales increase by 100 [.10 x 1000] *records.*

Interpreting the regression coefficient (slope, b_1)

regression coefficient, slope (b_1)

t statistic for b_1

p-value for the t statistic associated w/b_1

Model Coefficients - Sales Outcome

Predictor	Estimate	SE	t	p
Intercept	134.14	7.54	17.80	< .001
Predictor→ Adverts	0.10	0.01	9.98	< .001

What does it mean for *t* to be significant (p<.05) vs. not significant (p≥.05)?

- Significant *t*'s mean that the predictor is significantly related to the outcome.
- Non-significant *t*'s mean that the predictor is unrelated to the outcome.

Outline for Ch. 13 – Linear Regression

1. How to run linear regression using JAMOVI
2. How to interpret JAMOVI output
3. **How to write up results using APA style**

Reporting regression results using APA style

Results

Advertising budget was a significant positive predictor of record sales, $b = .10$, $t(198) = 9.98$, $p < .001$. For every £1000 increase in ad budget, sales increase by 100 records. Thirty-three percent of the variance in record sales is explained by ad budget, a large effect.

*For simple regression, use t(**N – 2**), where N is the sample size (shown on earlier slide as N = 200 record releases).*

Linear Regression

Model Fit Measures

Model	R	R^2
1	0.58	0.33

Model Coefficients - Sales

Predictor	Estimate	SE	t	p
Intercept	134.14	7.54	17.80	< .001
Adverts	0.10	0.01	9.98	< .001

Study Example two – Pubs and Death

A researcher investigated whether the number of pubs in a given region of London was related to how many deaths occurred in that region. *N* = 8 (8 regions of London)

Linear Regression

Model Fit Measures

Model	R	R^2
1	0.81	0.65

Model Coefficients - mortality

Predictor	Estimate	SE	t	p
Intercept	3351.96	781.24	4.29	0.005
pubs	14.34	4.30	3.33	0.016

Predictor (X): **number of pubs** Outcome (Y): **number of deaths**

Is there a significant relationship between the predictor and outcome? How do you know? **Yes, *p* = .016, which is less than .05 (the traditional significance level)**

Regression equation: $\hat{Y}$ = **3351.96 + 14.34X**

Results

Number of pubs in region of London was a significant positive predictor of number of deaths in that region (b = 14.34, $t(6)$ = 3.33, p = .016). For every one additional pub in a region, number of deaths increase by 14.34. Sixty-five percent of the variance in deaths can be explained by the number of pubs in the region, a very large effect.