Optimizing Advertising Budget Allocation

Kendra Mosholder

# Introduction

This analysis addresses a common business challenge: how to allocate a fixed advertising budget across multiple channels to maximize sales.  
Using a dataset of advertising spend on TV, radio, and newspapers, I applied regression models and budget simulations to evaluate each channel’s effectiveness.

The goal was to identify the most profitable allocation strategy while highlighting channel interactions, potential diminishing returns, and the limits of the available data.

# Packages

* tidyverse (dplyr, ggplot2, tidyr, readr, tibble)
* viridis (color-blind friendly palettes)

# Data

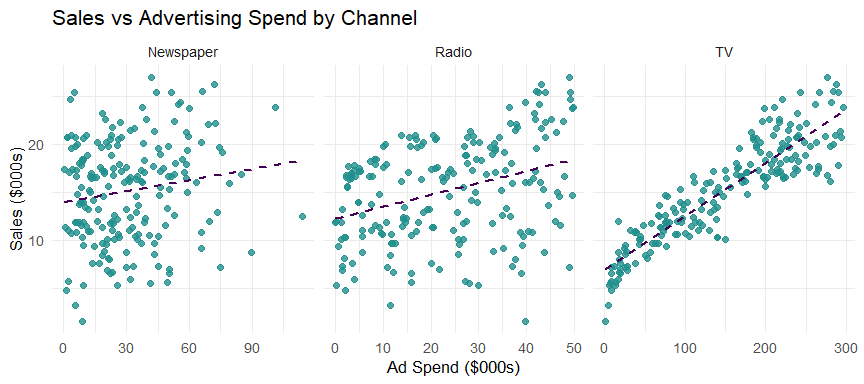
## [1] 200 4

## tv radio newspaper sales   
## Min. : 0.70 Min. : 0.000 Min. : 0.30 Min. : 1.60   
## 1st Qu.: 74.38 1st Qu.: 9.975 1st Qu.: 12.75 1st Qu.:11.00   
## Median :149.75 Median :22.900 Median : 25.75 Median :16.00   
## Mean :147.04 Mean :23.264 Mean : 30.55 Mean :15.13   
## 3rd Qu.:218.82 3rd Qu.:36.525 3rd Qu.: 45.10 3rd Qu.:19.05   
## Max. :296.40 Max. :49.600 Max. :114.00 Max. :27.00

## Insights

* 200 observations and 4 variables.
* TV spend up to ~296k, Radio up to ~50k, Newspaper up to ~114k, Sales up to ~27k.
* Data is clean and ready for analysis.

# Exploratory Data Analysis



## Insights

* TV has the strongest positive relationship with sales.
* Radio shows a moderate positive relationship.
* Newspaper shows little to no relationship with sales.

# Baseline Regression

##   
## Call:  
## lm(formula = sales ~ tv + radio + newspaper, data = ads)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -7.3034 -0.8244 -0.0008 0.8976 3.7473   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.6251241 0.3075012 15.041 <2e-16 \*\*\*  
## tv 0.0544458 0.0013752 39.592 <2e-16 \*\*\*  
## radio 0.1070012 0.0084896 12.604 <2e-16 \*\*\*  
## newspaper 0.0003357 0.0057881 0.058 0.954   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.662 on 196 degrees of freedom  
## Multiple R-squared: 0.9026, Adjusted R-squared: 0.9011   
## F-statistic: 605.4 on 3 and 196 DF, p-value: < 2.2e-16

## Insights

* TV and Radio are statistically significant predictors of sales.
* Newspaper is not significant.
* The model explains about 90% of the variation in sales (Adjusted R² ≈ 0.90).

# Interaction Model

##   
## Call:  
## lm(formula = sales ~ tv \* radio + newspaper, data = ads)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -6.2685 -0.8765 -0.0480 0.9339 3.6521   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 6.172e+00 4.192e-01 14.722 < 2e-16 \*\*\*  
## tv 4.355e-02 2.498e-03 17.433 < 2e-16 \*\*\*  
## radio 4.146e-02 1.513e-02 2.740 0.00671 \*\*   
## newspaper 1.349e-03 5.454e-03 0.247 0.80492   
## tv:radio 4.439e-04 8.699e-05 5.103 7.91e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.565 on 195 degrees of freedom  
## Multiple R-squared: 0.9141, Adjusted R-squared: 0.9123   
## F-statistic: 518.6 on 4 and 195 DF, p-value: < 2.2e-16

## Insights

* The TV × Radio interaction is statistically significant.
* Combining channels produces stronger results than using them alone.
* Adjusted R² ≈ 0.91 (slight improvement over baseline).

# Curved Model (Diminishing Returns Test)

##   
## Call:  
## lm(formula = sales ~ tv + I(tv^2) + radio + I(radio^2) + tv:radio,   
## data = ads)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.0642 -0.7938 -0.0079 0.7480 3.4044   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 5.153e+00 4.499e-01 11.452 < 2e-16 \*\*\*  
## tv 7.554e-02 4.882e-03 15.473 < 2e-16 \*\*\*  
## I(tv^2) -1.085e-04 1.506e-05 -7.205 1.25e-11 \*\*\*  
## radio -3.354e-02 2.712e-02 -1.237 0.21767   
## I(radio^2) 1.770e-03 5.149e-04 3.437 0.00072 \*\*\*  
## tv:radio 4.145e-04 7.594e-05 5.459 1.45e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.363 on 194 degrees of freedom  
## Multiple R-squared: 0.9351, Adjusted R-squared: 0.9335   
## F-statistic: 559.3 on 5 and 194 DF, p-value: < 2.2e-16

## Insights

* The negative TV² term suggests possible diminishing returns.
* The dataset lacks higher budget levels to confirm this pattern.
* Within the observed range, sales keep increasing with spend.
* Adjusted R² ≈ 0.93 (best fit so far).

# Budget Allocation Simulation

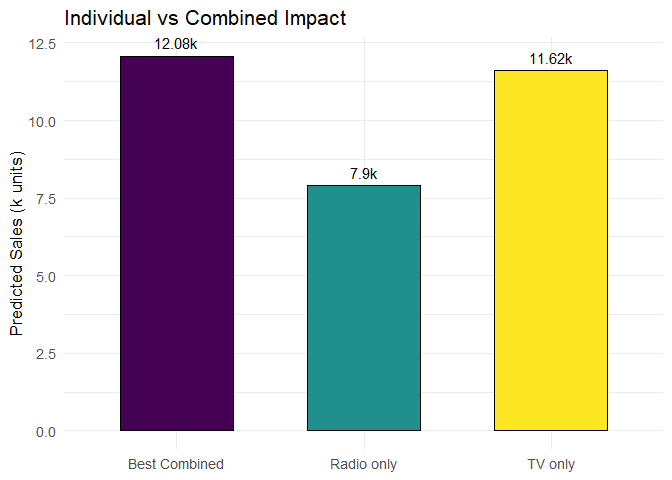
## tv radio pred\_sales  
## 1 55 45 12.07971

## Insights

* Best allocation ≈ TV 55k / Radio 45k.
* Predicted sales ≈ 12.1k.
* A balanced mix outperforms single-channel spend.

# Visualizations

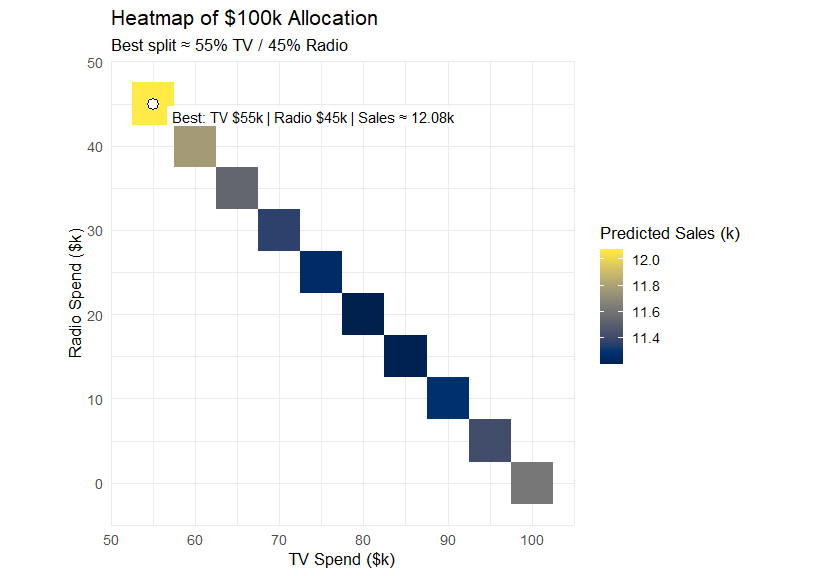
## Individual vs Combined Impact



## Insights

* TV only ≈ 11.6k sales.
* Radio only (capped at 50k) ≈ 7.9k.
* Best combined ≈ 12.1k sales.

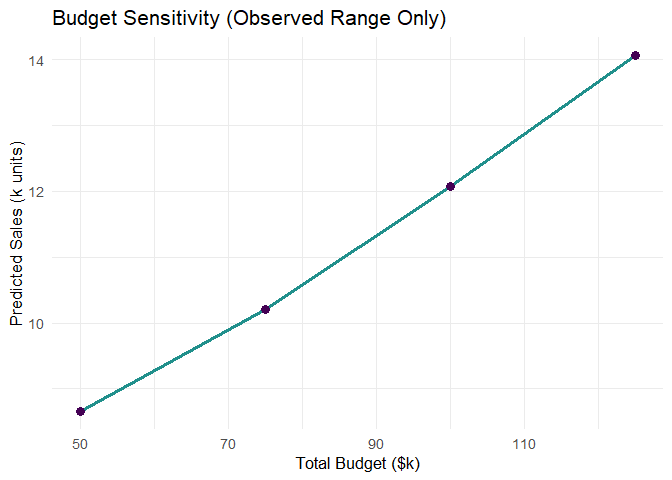
# Heatmap of $100k Allocation (Self‑Contained)



## Insights

* Heatmap confirms best allocation ≈ 55/45 split.
* Newspaper excluded (not predictive).

# Budget Sensitivity (Observed Range)



## Insights

* From $50k to $125k, predicted sales rise steadily.
* No saturation point is visible in this dataset.
* More data is needed at higher budgets to test for diminishing returns.

# Recommendations

* Allocate ~$100k across TV and Radio with an optimal split near **55% TV / 45% Radio**.
* Do not allocate to Newspaper (no measurable effect in this dataset).
* If budgets increase beyond $100k, collect additional data to test for diminishing returns.
* Incorporate digital channels in future analyses for completeness.

# Key Insights

* I framed a business question (how to spend a fixed ad budget) as a data problem.
* I used regression and simulation to compare scenarios and find an optimal mix.
* Multichannel (TV + Radio) outperforms single-channel within the observed range.
* I recognized dataset limits and proposed next steps.

# Reproducibility

This analysis was generated in R using the following session information:

## tidyverse viridis   
## "2.0.0" "0.6.5"