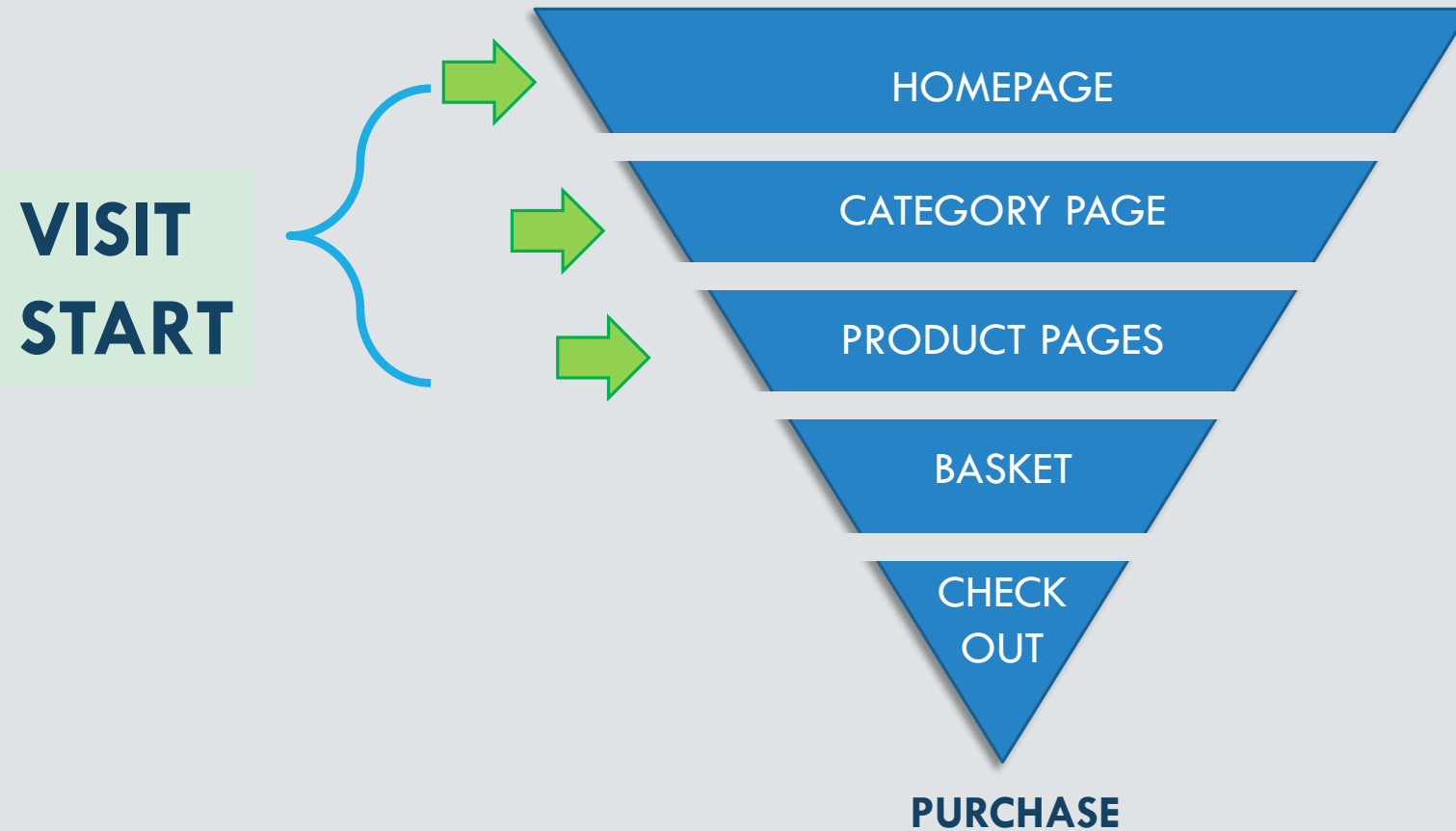




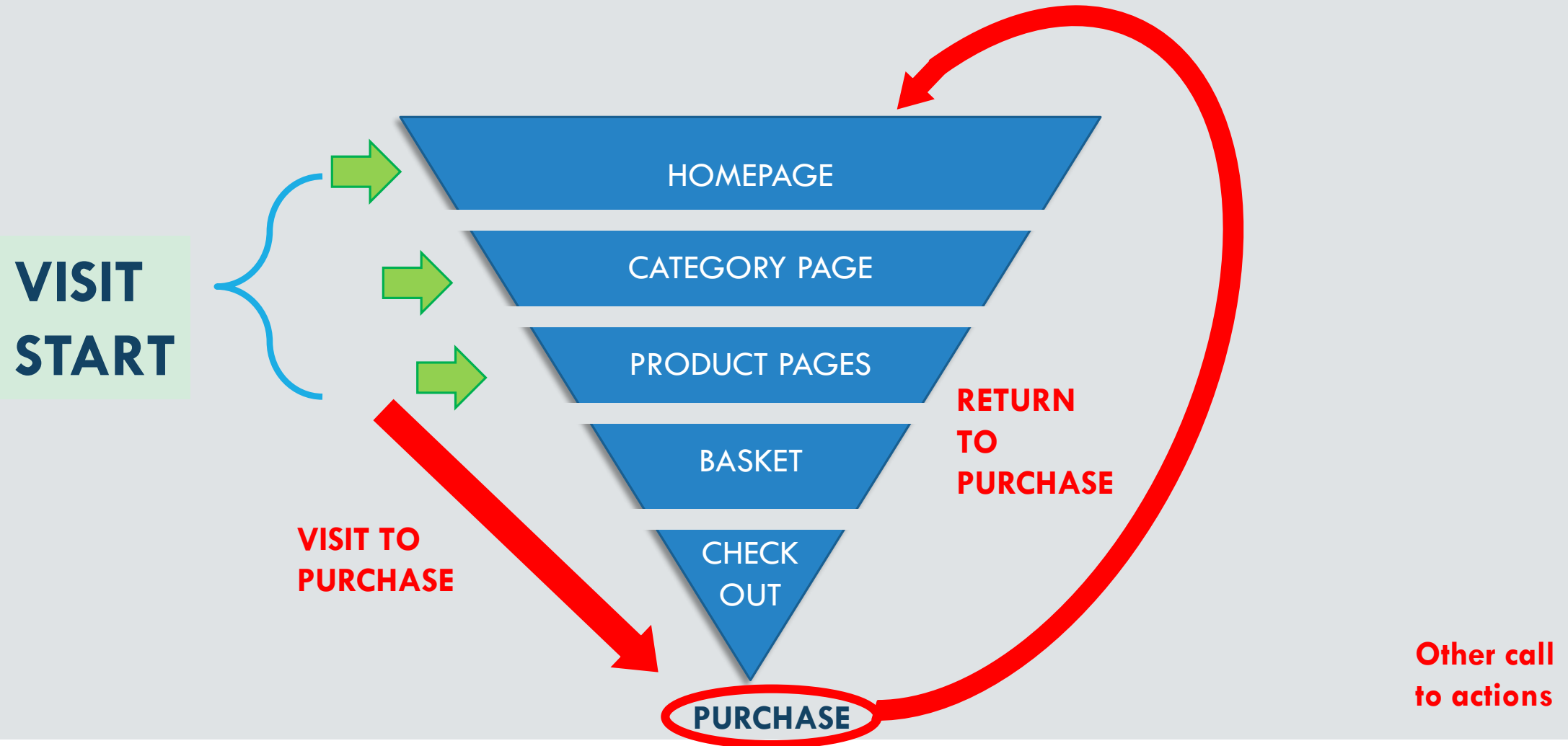
# USING R TO UNDERSTAND REVENUE BENEFIT FOR YOUR ONLINE BUSINESS

Catie Gamble-Miller  
Senior Data Scientist  
Marks and Spencer .com

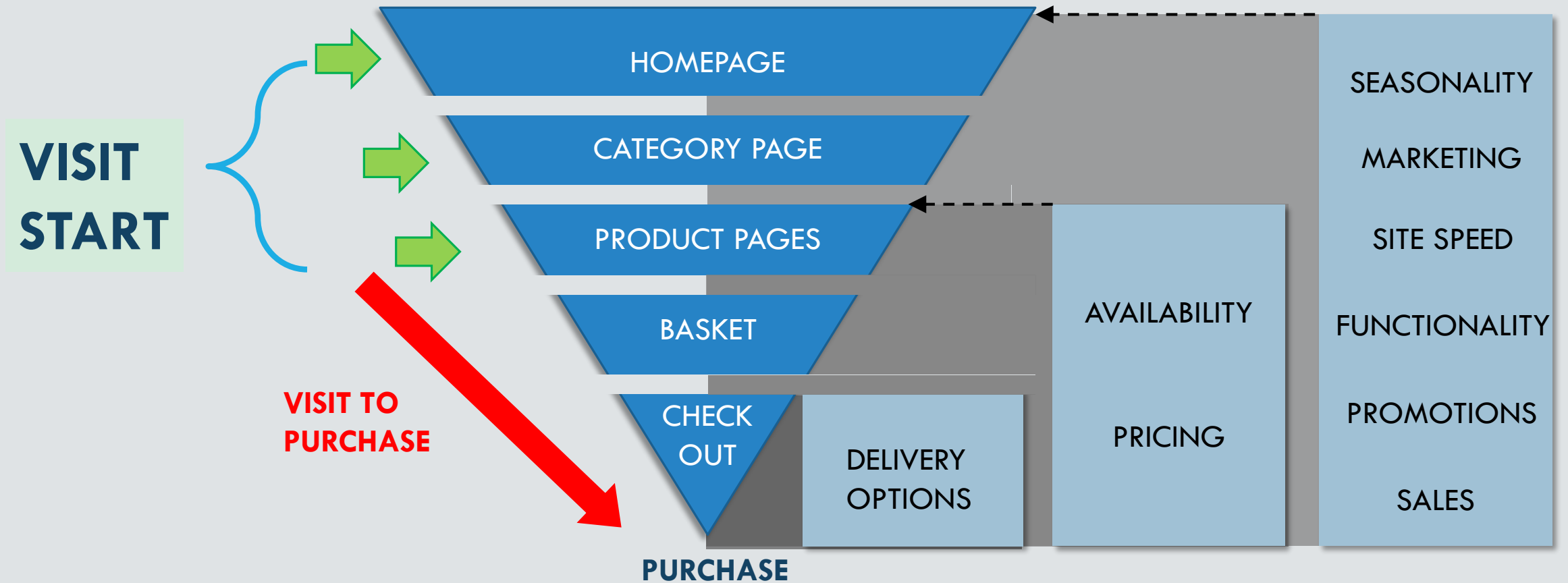
# ONLINE RETAIL BASICS



# ONLINE RETAIL BASICS



# ONLINE RETAIL BASICS



# PROBLEMS FACED

New feature added to site during promotional period  
We know the new feature drove higher conversion or customer engagement  
**But by how much?**

Potential new initiative or strategy will bring additional revenue  
But is set to degrade performance/delivery/etc...  
**How much additional revenue will be offset?**

Scenario test several options for the same initiative  
**Which option is the most profitable/engages the customer the most?**

**and...Feature/Initiative cannot be measured by AB test**

# EXAMPLE IN R — THE PROBLEM

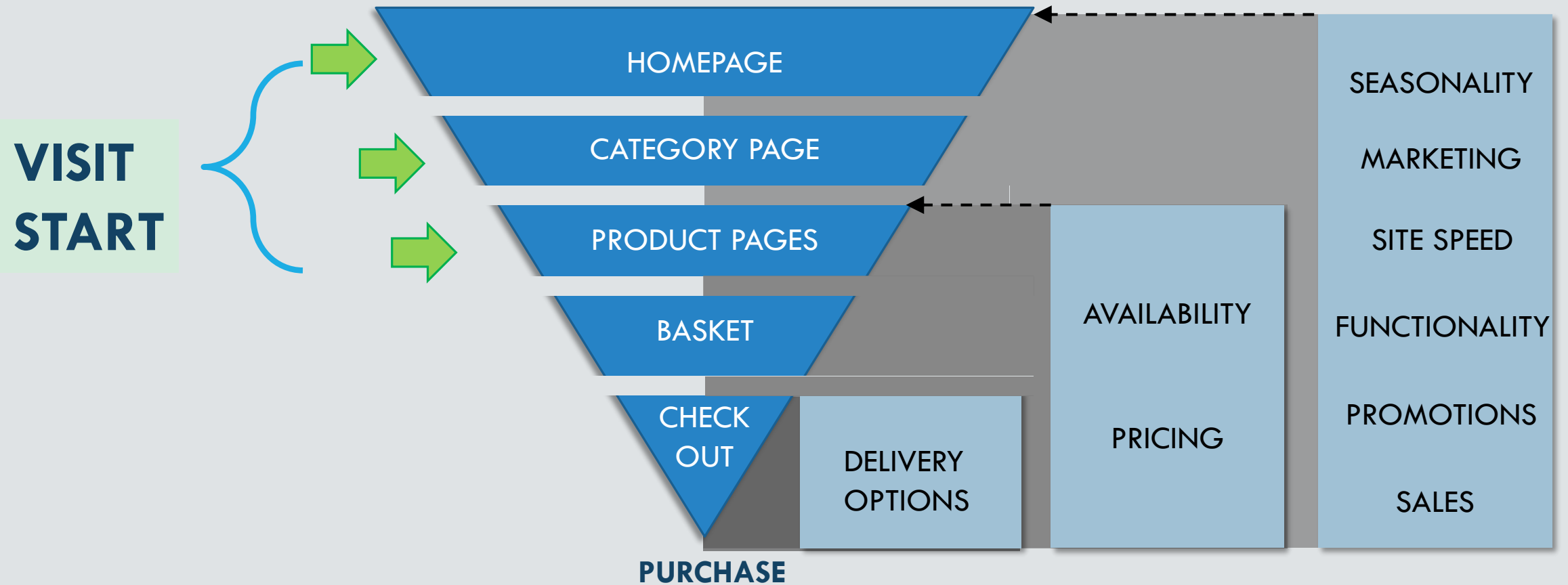
New feature added to site during promotional period  
We know the new feature drove higher conversion or customer engagement  
**But by how much?**

Potential new initiative or strategy will bring additional revenue  
But is set to degrade performance/delivery/etc...  
**How much additional revenue will be offset?**

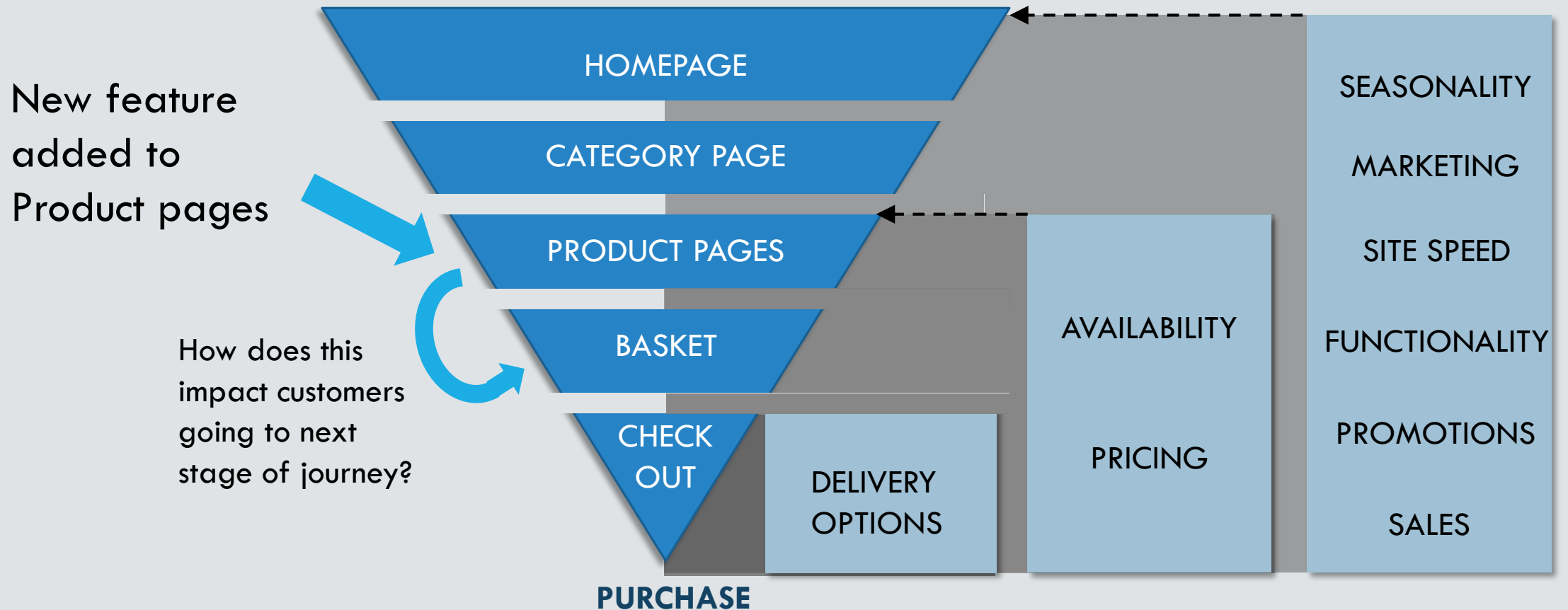
Scenario test several options for the same initiative  
**Which option is the most profitable/engages the customer the most?**

**and...Feature/Initiative cannot be measured by AB test**

# ONLINE RETAIL BASICS



# EXAMPLE IN R – THE PROBLEM

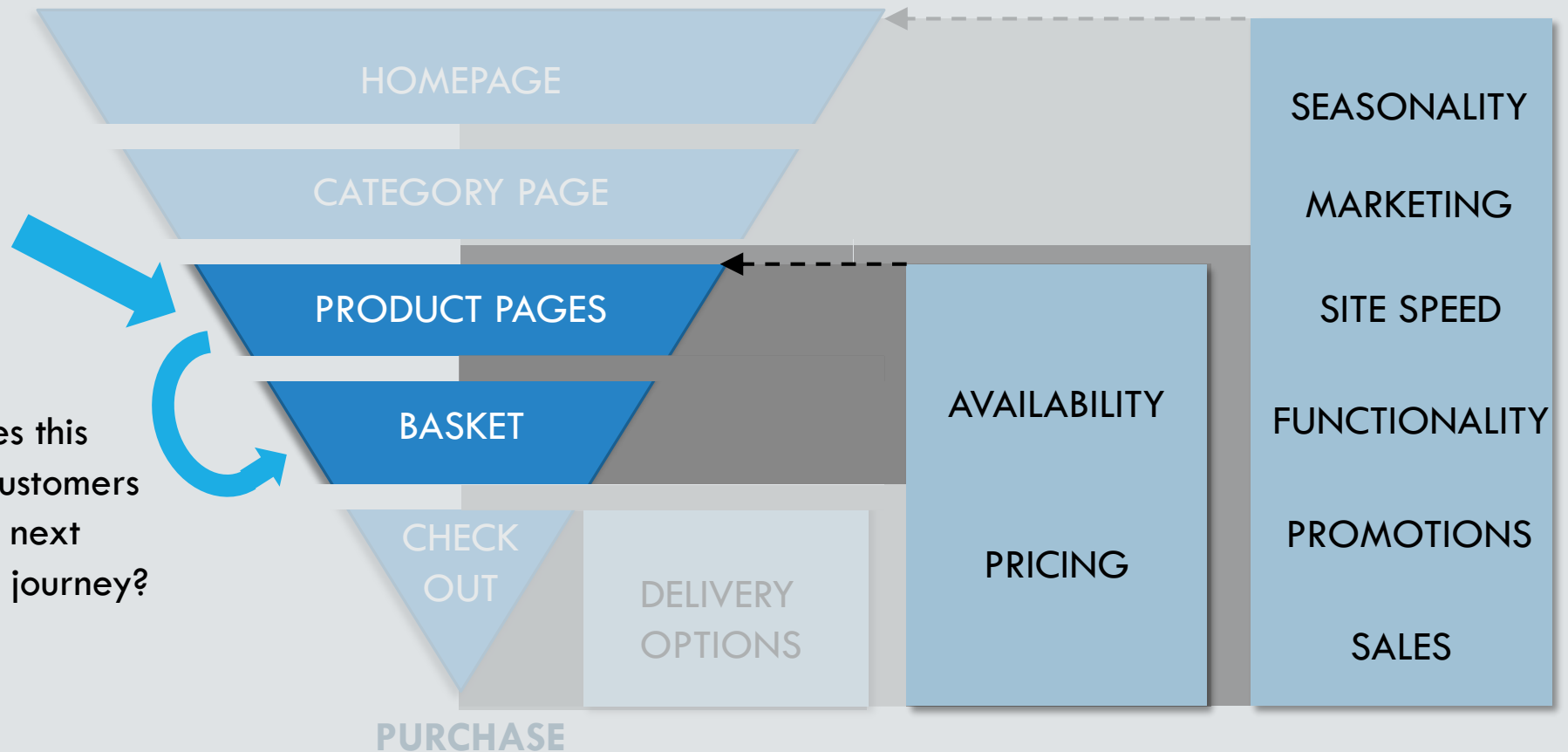




# EXAMPLE IN R – THE PROBLEM

New feature added to Product pages

How does this impact customers going to next stage of journey?



# EXAMPLE IN R – REGRESSION

Product Page to  
Add to Basket  
Conversion Rate  
%



=

Intercept

+



+

Impact of  
New feature  
added to  
Product pages

# EXAMPLE IN R – REGRESSION – STEPS

## 1 Format data + exploratory analysis

date	device	ATB cr	
			...
			...

## 2

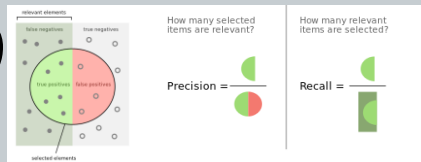
Build model.

```
> model<-glm(pdp_atb_rate~seasonality_data
+ +marketing_data+page_load_data
+ +promo_sales_data+availability_data
+ +pricing_data+...
+ ,family=binomial(link=logit),data)
```

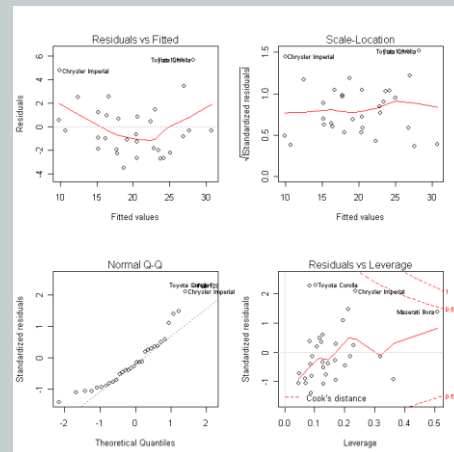
Consider interaction

variables `page load data*promo sales data`

## 3 Model diagnostics




$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$



## 4

Financial uplift

$$\text{Impact on ATB Rate} \times \text{Visits} \times \text{Conversion Rate} \times \text{AOV} = \text{Potential Revenue Impact}$$


# WHAT TO MODEL?

## Marketing

### Propensity Modelling

```
> glm(, , family=  
+ binomial(link=logit)) > rpart()  
> randomForest()
```

### Marketing/Visit Forecasting

```
> glm(, , family=poisson)
```

**VISIT  
START**

## Funnel KPI modelling

HOMEPAGE

Exit rate  
predictions

```
> glm()  
> lm()
```

CATEGORY PAGE

PRODUCT PAGES

Add to basket

```
> glm(, , family=  
+ binomial(link=logit))  
> rpart()  
> randomForest()
```

BASKET

CHECK  
OUT

Checkout to purchase > rpart()

```
> glm(, , family=  
+ binomial(link=logit))  
> lm() > randomForest()
```

**PURCHASE**

Customer  
segmentation

```
> kmeans()  
> hclust()
```

Control groups

```
> rpart()  
> randomForest()  
> glm(, , family=  
+ binomial(link=logit))
```