



GUS' GOOD SPREADSHEET GUIDE

By: Gus Lipkin, Spreadsheet Enthusiast



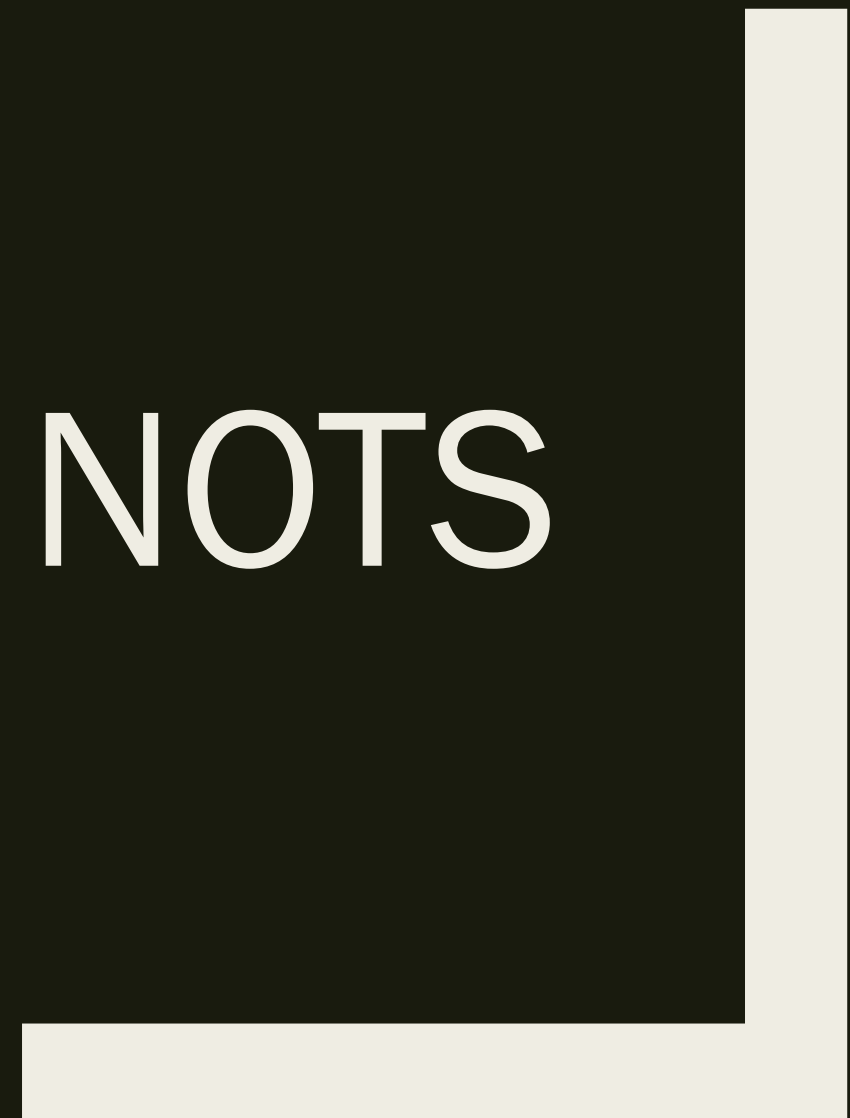
Why does this slideshow exist?

- Spreadsheets are an efficient tool for communicating large amounts of information
- Spreadsheets can be incredibly customized to the user's needs
- Customized spreadsheets are very difficult to write programs for

Thus,

- Generating new spreadsheets becomes a much longer and more complicated process than necessary

DO AND DO NOTS



Symbols

Do not use symbols or spaces

Food_Type	UPC #	Type
Apple	F-27753	Fruit
Banana	F-226262	Fruit
Carrot	V-227768	Veg.

Do make everything one word without symbols

FoodType	UPC	Type
Apple	F27753	Fruit
Banana	F226262	Fruit
Carrot	V227768	Vegetable

- Special characters are difficult to decipher and may look the same but not hold the same value

Cell Commenting

Do not use cell commenting

FoodType	UPC	
Apple	F27753	Bad
Banana	F226262	Okay
Carrot	V227768	Good

Do use more columns

FoodType	UPC	Notes
Apple	F27753	Bad
Banana	F226262	Okay
Carrot	V227768	Good

- Cell commenting cannot be imported

DO AND DO NOTS

Naming



Sheet and Column Names

- Should be clear and concise
- All columns and sheets should be named
- Column and sheet names should never start with a number
- Column names should be unique in each workbook for sheet names and unique to each sheet for column names
- Should contain only letters and numbers

Column Names 1

Do not use different column names in different spreadsheets

FoodType	UPC	Type	
App	Food	ProductCode	FruitOrVeg
Ban	Apple	F27753	Fruit
Car	Banana	F226262	Fruit
	Carrot	V227768	Vegetable

Do use the same column names in different spreadsheets

FoodType	UPC	Type	
App	FoodType	UPC	Type
Ban	Apple	F27753	Fruit
Car	Banana	F226262	Fruit
	Carrot	V227768	Vegetable

- Inconsistent column names cause difficulties when matching data between spreadsheets

Column Names 2

Do not use spaces and/or inconsistent capitalization

FOODTYPE	upc	Type
Apple	F27753	Fruit
Banana	F226262	Fruit
Carrot	V227768	Vegetable

Do use TitleCase or camelCase with no spaces

FoodType	UPC	Type
Apple	F27753	Fruit
Banana	F226262	Fruit
Carrot	V227768	Vegetable

- Inconsistent grammar makes it difficult to import data
- TitleCase capitalizes the first letter of each word in the string
 - ThisIsTitleCase
- camelCase capitalizes the first letter of each word in the string except the first
 - thisIsCamelCase

Column Names 3a

Do not put anything besides column names in the first row

Gus' Grocery Store		
FoodType	UPC	Type
Apple	F27753	Fruit
Banana	F226262	Fruit

Do put column names in the first row

FoodType	UPC	Type
Apple	F27753	Fruit
Banana	F226262	Fruit
Carrot	V227768	Vegetable

- Column names cannot be imported if they are not the first row

Sheet Names

'Do Not' Examples

FoodType	UPC	Type
Apple	F27753	Fruit
Banana	F226262	Fruit
Carrot	V227768	Vegetable

Produce sheet 2 baked_goods

'Do' Examples

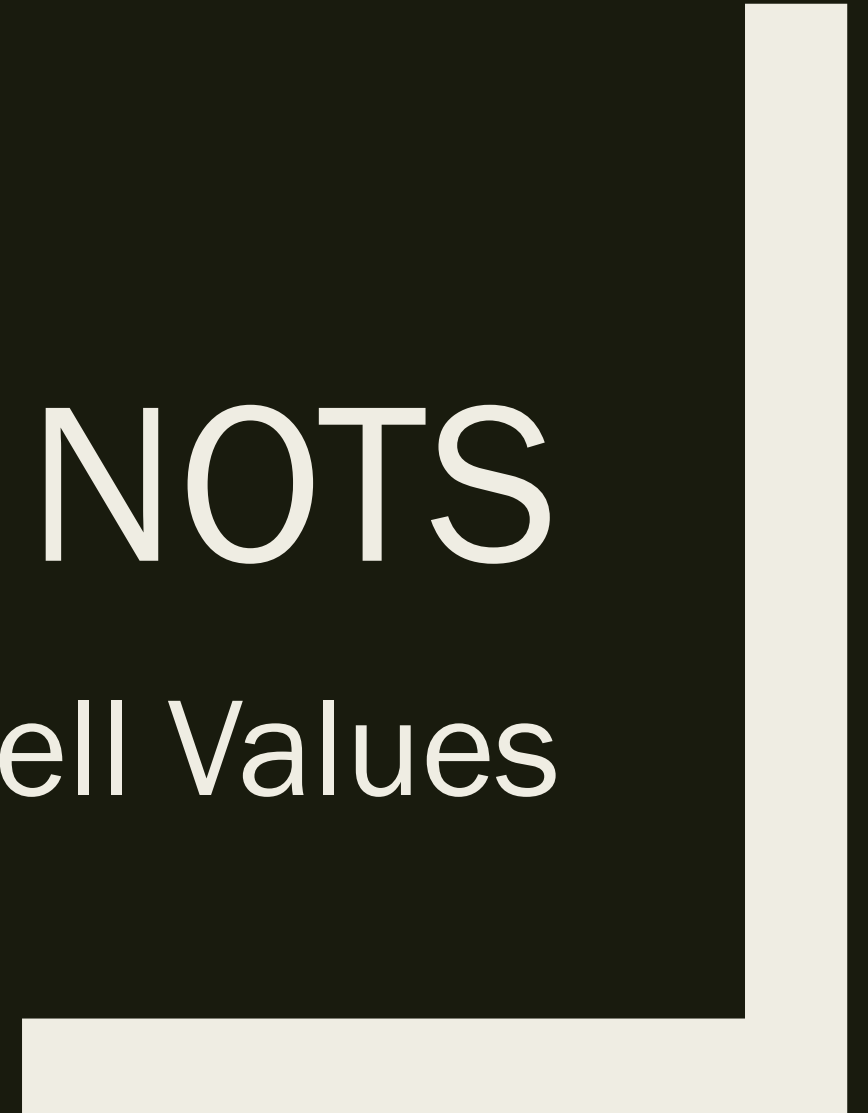
FoodType	UPC	Type
Apple	F27753	Fruit
Banana	F226262	Fruit
Carrot	V227768	Vegetable

Produce Sheet2 BakedGoods

- Sheet names should follow the same conventions as column names

DO AND DO NOTS

Cell Values



Column Values

Do not use inconsistent formatting

Food type	UPC #	type
apple	27753	fruit
BANANA	F226262	FRUIT
Carrot	V-227768	Vegetable

Do use consistent formatting

FoodType	UPC	Type
Apple	F27753	Fruit
Banana	F226262	Fruit
Carrot	V227768	Vegetable

- It takes time to discern true differences between inconsistent formats

Boolean Columns

Do not use arbitrary designations

FoodType	UPC	IsFruit
Apple	F27753	Yes
Banana	F226262	yes
Carrot	V227768	no

Do use TRUE and FALSE

FoodType	UPC	IsFruit
Apple	F27753	TRUE
Banana	F226262	TRUE
Carrot	V227768	FALSE

- A Boolean value is a value that has only two states
- Boolean values should be either TRUE or FALSE
 - While 'Yes' and 'No' are easy to understand, it is easy to accidentally type 'yes' or 'NO' which are not the same
 - Similarly, 1 and 0 or x and o are not immediately clear what data type is in that column

Intentionally Blank Cells 1 (Numbers)

Do not leave blank cells

FoodType	UPC	DaysSinceRipe
Apple	F27753	10
Banana	F226262	0
Carrot	V227768	

Do use -1 when intentionally left blank

FoodType	UPC	DaysSinceRipe
A4	F27753	10
F1	F226262	0
F1	V227768	-1

- A number column should never have a negative value, this means that -1 can be reserved for intentionally blank cells
- -1 can also mean infinity. This is because a number not existing is the same as any number existing
- When a cell is left blank, its value cannot be assumed
 - This can cause issues down the road

Intentionally Blank Cells 2 (Text)

Do not leave blank cells

FoodType	UPC	Variety
Apple	F27753	Abram
Banana	F226262	Blue Java
Carrot	V227768	

Do use FALSE when intentionally left blank

FoodType	UPC	Variety
Apple	F27753	Abram
Banana	F226262	Blue Java
Carrot	V227768	FALSE

- Even though it is a text column, FALSE can be properly interpreted as not having value
- When a cell is left blank, its value cannot be assumed
 - This can cause issues down the road

Intentionally Blank Cells 3 (Booleans)

Do not leave blank cells

FoodType	UPC	InStock
Apple	F27753	TRUE
Banana	F226262	FALSE
Carrot	V227768	

Do use FALSE when intentionally left blank

PlaneType	UPC	InStock
Apple	F27753	TRUE
Banana	F226262	FALSE
Carrot	V227768	FALSE

- Because a Boolean is either TRUE or FALSE, it is safe to assume that if it is not TRUE, then it is FALSE
 - This is the only time it is safe to assume a blank cell's value

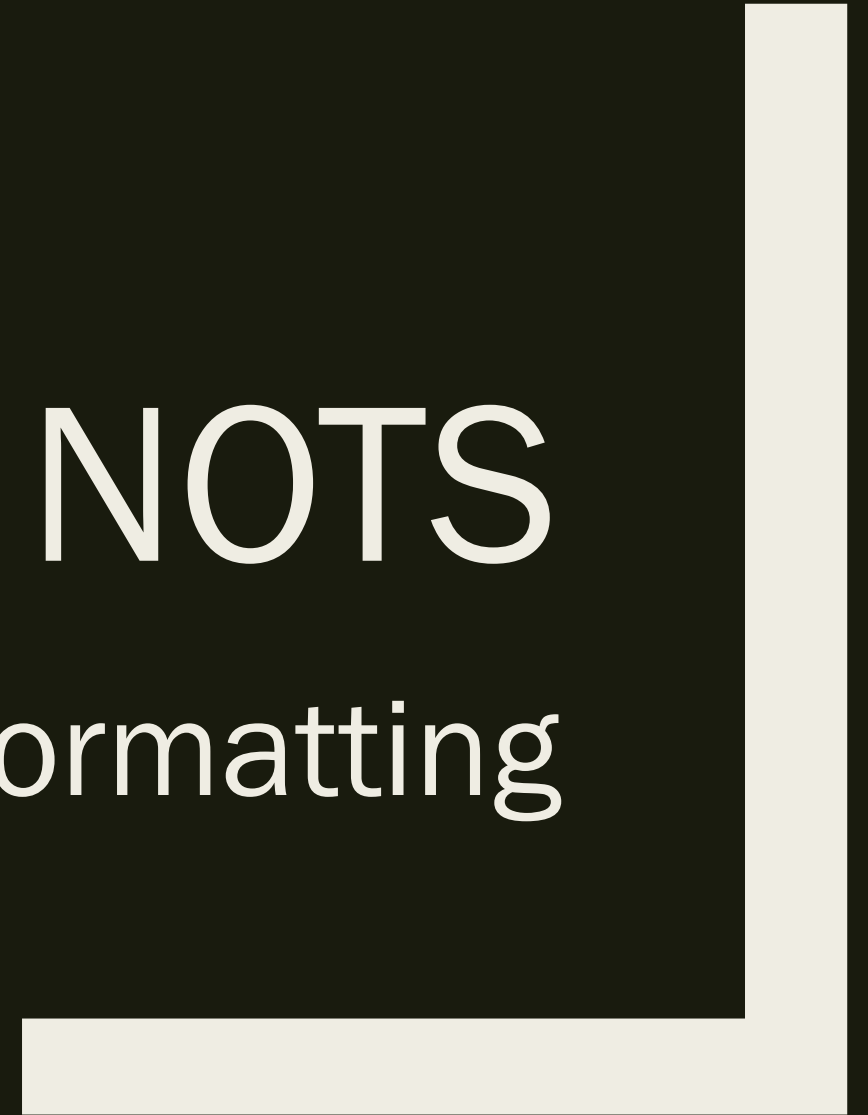
Intentionally Blank Cells 4

- If none of these may seem quite right for your use case...

Ultimately it is up to you to make the best decision, but it should always be clear and obvious which cells are intentionally blank

DO AND DO NOTS

Formatting



Merged Cells

Do not merge cells

FoodType	OldUPC	NewUPC
Apple	F27753	
Banana	F226262	F226262BJ
		F226262C

Do repeat data

FoodType	OldUPC	NewUPC
Apple	F27753	F27753
Banana	F226262	F226262BJ
Banana	F226262	F226262C

- Merged cells are difficult to work with and cannot be easily split into the proper format

Cell Formatting

Do not use cell highlighting

FoodType	UPC
Apple	F27753
Banana	F226262
Carrot	V227768

Do use more columns

FoodType	UPC	StockLevel
Apple	F27753	Out
Banana	F226262	Low
Carrot	V227768	Good

- Cell formatting cannot be imported. This includes but is not limited to
 - Highlighting
 - Borders

Text Formatting

Do not use text formatting

FoodType	UPC
Apple	<u>F27753</u>
Banana	F226262
Carrot	V227768

Do use more columns

FoodType	UPC	StockLevel
Apple	F27753	Out
Banana	F226262	Low
Carrot	V227768	Good

- Text formatting cannot be imported. This includes but is not limited to
 - *Italics*, **Bold**, Underlines, ~~Strikethroughs~~
 - Font colors

Data Formatting 1

Do not make new columns to hold a list

FoodType	Variety1	Variety2
Apple	Abram	Bailey
Banana	Blue Java	Cavendish

Do repeat the primary column

FoodType	Variety
Apple	Abram
Apple	Bailey
Banana	Blue Java
Banana	Cavendish

- Columns should hold variable values by row

Data Formatting 2

Do not use multiple tables in one sheet

AppleType	UPC	
Abram	F27753A	
	BananaType	UPC
	Blue Java	F226262BJ

Do split different tables into different sheets

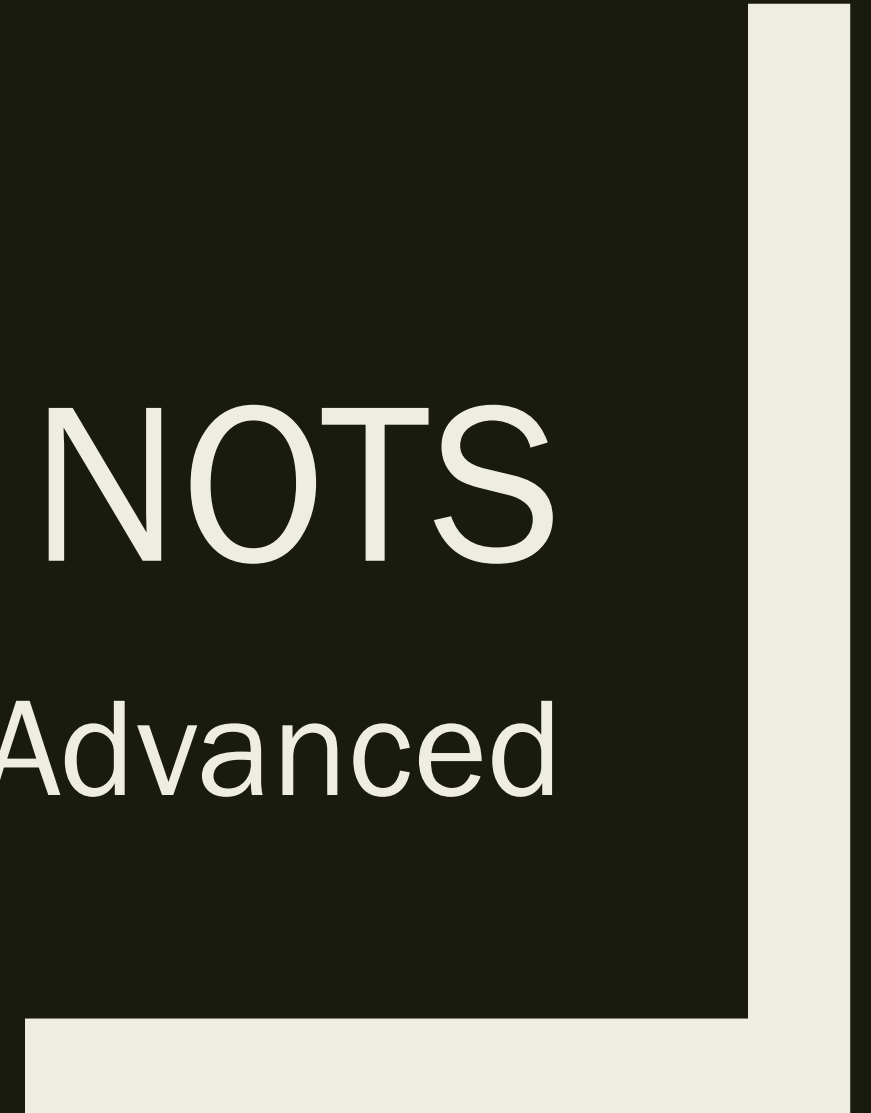
Variety	UPC
Abram	F27753A
Bailey	F27753B
Cabashea	F27753C

AppleTypes BananaTypes

- Multiple tables in one sheet are unusable

DO AND DO NOTS

Advanced



Advanced Dos

- If data is split across multiple sheets, the first column should be the same across sheets
 - This is known as a primary key and is used to make sure the data is consistent across sheets
 - If a primary key does exist, the previous slides are a complete guide

- Footnote: This is a vast oversimplification and technically incorrect application of the term “primary key” but is suitable for our purposes here

Column Names 3b

If you have duplicate column names that require two header rows...

TestingStageOne		TestingStageTwo	
Pass	CheckedBy	Pass	CheckedBy
TRUE	AMY	FALSE	BOB
FALSE	AMY	TRUE	BOB

Use multiple rows as described in [Data Formatting 1](#)

TestingStage	Pass	CheckedBy
1	TRUE	AMY
2	FALSE	BOB
1	FALSE	AMY
2	TRUE	BOB

- This is only possible if you can have duplicate rows with your primary key and is less human readable

Column Names 3c

If you have duplicate columns that require two header rows and cannot duplicate rows

TestingStageOne		TestingStageTwo	
Pass	CheckedBy	Pass	CheckedBy
TRUE	AMY	FALSE	BOB
FALSE	AMY	TRUE	BOB

Use a period to distinguish them

TestOne.Pass	TestOne.CheckedBy	TestTwo.Pass	TestTwo.CheckedBy
TRUE	AMY	FALSE	BOB
FALSE	AMY	TRUE	BOB

- This should only be done if you cannot use multiple rows as described in [Column Names 3b](#)
- This is more human readable than [Column Names 3b](#)
- This is the only time it is acceptable to use a period in a column name

Data Formatting 3

If a cell needs to hold more than one piece of information and a primary key doesn't exist

FoodType	UPC1	UPC2
Apple	F27753A	F27753B
Banana	F226262	

Combine [Data Formatting 2](#) and [Cell Formatting](#)

Banana	Apple
F226262	F27753A

Banana	Apple
Good	Low
	Out

UPC

StockLevel

- Transpose the data then split it into two sheets
- The structure of the data must be the same between sheets

Part Numbers

Do not list multiple part numbers in the same cell

UPC	FoodType
F27753	Apple
F226262BJ/F226262C	Banana
V227768A or V227768B or V227768C	Carrot

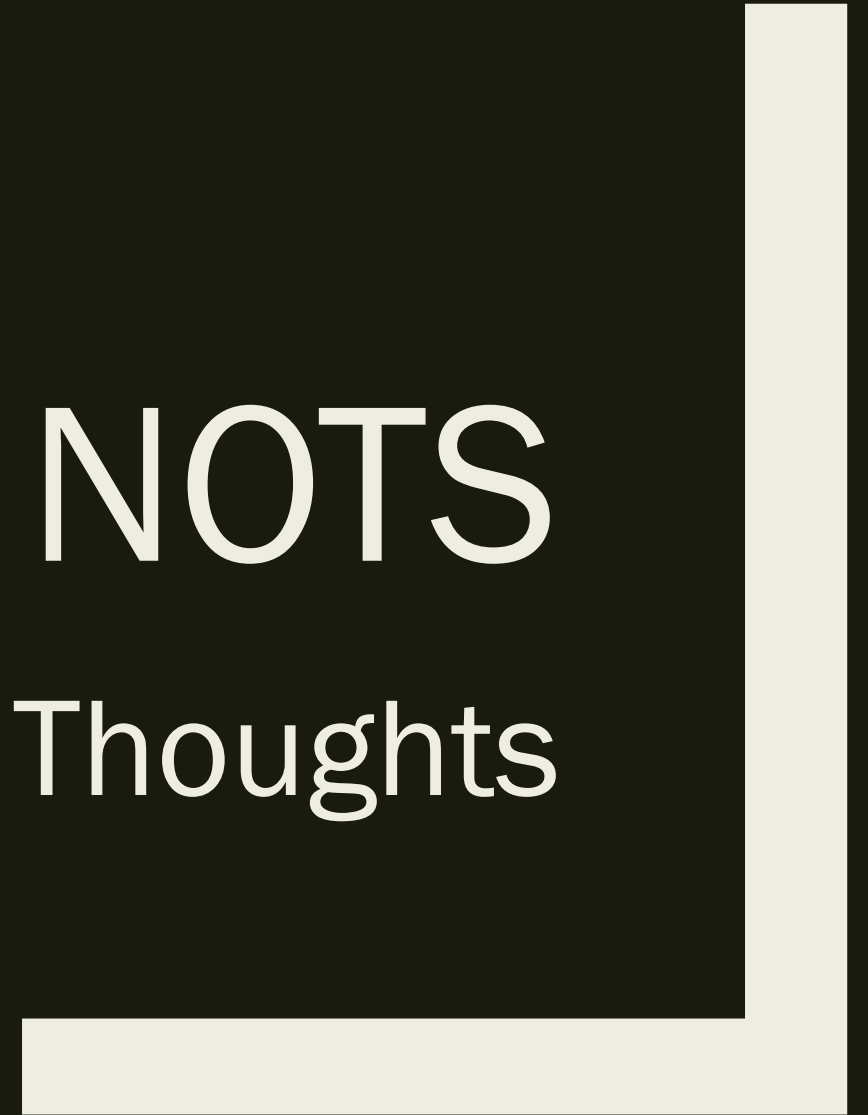
Do use more columns and duplicate primary part numbers

BaseUPC	VarietyUPC	FoodType
F27753	F27753	Apple
F226262	F226262BJ	Banana
F226262	F226262C	Banana
V227768	V227768A	Carrot
V227768	V227768B	Carrot
V227768	V227768C	Carrot

- Choose one part number to be the primary part number
- List all alternate part numbers in their own cell in a separate column
- Duplicate primary part numbers where necessary

DO AND DO NOTS

Parting Thoughts



Parting Thoughts

- As lengthy as this presentation may be, it is by no means exhaustive. If you have any questions about anything that is or is not covered here, please ask.