

# Check Digits

It's the little things that tax us and put us on  
the rack.

You can sit upon a mountain but not upon a  
tack

***-Anonymous***

# Identifiers

- Social Security Number
- Vehicle Identification Number
- NM Drivers License Number
- UPC (Bar Codes)
- ISBN
- U. S. Post Office Money Orders
- Zip Codes (12 digit)
- Credit Card Numbers

# Common Features

- Except for Social Security Numbers, all have check digits.
- Except for NM Drivers Licenses, all contain coded information.

# Why use check digits?

- Machines can make errors.
  - ❖ Transmission errors.
- People can make errors.
  - ❖ SSN on scantrons.

# Types of Errors

Type	Form	Rel. Freq.	Correct	Incorrect
Single Digit	$a \rightarrow b$	79.1%	<u>132</u>	<u>182</u>
Trans.	$ab \rightarrow ba$	10.2%	<u>5431</u>	<u>5341</u>
Jump	$abc \rightarrow cba$	0.8%	<u>76913</u>	<u>71963</u>
Twin	$aa \rightarrow bb$	0.5%	<u>1776</u>	<u>1886</u>
Phonetic	$a0 \rightarrow 1a$	0.5%	<u>1140</u>	<u>1114</u>
Jump Twin	$aca \rightarrow bcb$	0.3%	<u>25859</u>	<u>23839</u>

# Zip Code (12 digit)

- The bar code is intended to be read by a scanner.
- The ZIP code determined by the scanner is used in automated mail sorting facilities to route pieces of mail.
- Each digit is represented by five vertical bars. Two of which are taller than the others.

# Zip Code (12 digits)

- Form: gssll-rrrr-hhc
  - ❖ g, represents one of ten geographical areas.
  - ❖ ss, identifies a sectional center
  - ❖ ll, represents the local town or local post office.
  - ❖ The next four locate the route
  - ❖ The next two the house
  - ❖ The last is a check digit

# Zip Code Check Digit

- Check
  - ❖ Add all twelve digits. If the number is correct, the sum will be an even multiple of ten.
- Calculation
  - ❖ Add the first eleven digits.
  - ❖ The check digit is the number needed to make the sum an even multiple of ten.

# Zip Code Analysis

- Advantages
  - ❖ This will catch all single digit errors.
  - ❖ It is easy to calculate
- Disadvantages
  - ❖ It will not catch transposition errors

# Zip Code Example

- **88003-8001-011**

$$\blacklozenge 8 + 8 + 0 + 0 + 3 + 8 + 0 + 0 + 1 + 0 + 1 + 1 = 30$$

- **88003-7001-011**

$$\blacklozenge 8 + 8 + 3 + 7 + 1 + 1 + 1 = 29$$

- **88003-8010-011**

$$\blacklozenge 8 + 8 + 0 + 0 + 3 + 8 + 0 + 1 + 0 + 0 + 1 + 1 = 30$$

# UPC (12 digit, Version A)

- The bar code is intended to be read by a scanner.
- It is divided into left and right halves.
- The codes for digits in the left half are mirror images of those in the right half.
  - ❖ This allows the code to be read either from right to left or vice versa.

# UPC (12 digit Version A)

- Form: P LLLLL RRRRR C
  - ❖ The first digit is the “product type.”
  - ❖ The next five form the remainder of the left half and include the code for the manufacturer.
  - ❖ The digits seven through eleven form the right half and represent the code for the product.
  - ❖ The final digit is the check digit.

# UPC Check Digit

- Check:
  - ❖ Add the odd numbered digits, multiply this by three, then add the even numbered digits. If the number is correct, the sum will be an even multiple of ten.
- Calculation:
  - ❖ Add the odd numbered digits of the base number, multiply this by three, add the even numbered digits. The check digit is the number that makes the sum an even multiple of ten.

# UPC Analysis

- Advantages:
  - ❖ This will catch all single digit errors.
  - ❖ This will catch most transposition errors.
  - ❖ It is easy to calculate.
- Disadvantages:
  - ❖ It will not catch transposition errors when the difference between the digits is five.
  - ❖ It will not catch Jump Transposition errors.

# UPC Example

- **0-78910-29403-1**

- ❖  $3(0 + 8 + 1 + 2 + 4 + 3) + (7 + 9 + 0 + 9 + 0 + 1)$

- ❖  $3(18) + 26 = 54 + 26 = 80$

- **0-78910-29143-1**

- ❖  $3(0 + 8 + 1 + 2 + 1 + 3) + (7 + 9 + 0 + 9 + 4 + 1)$

- ❖  $3(15) + 30 = 75$

- **0-78910-24903-1**

- ❖  $3(23) + 21 = 90$

# International Standard Book Number (ISBN)

- This is meant to uniquely identify books.
- It was not developed to be scanned.
- It is an international standard

# ISBN

- Form: L-PPPPP-BBBBBB-C
- The ISBN is shown in four groups of numbers. The first number identifies the country, or language group, the second the publisher, the third the book, and the fourth is a check digit.
- It is a ten digit number, the groups may have any length provided the total length of the number is ten digits.

# ISBN Check Digit

- Check:
  - ❖ Each digit has a “place” value. The first digit has a place value of ten, the next a nine, etc. until the tenth digit which has a place value of one.
  - ❖ Multiply each digit by its place value and add. The sum will be a multiple of eleven.
- Calculation:
  - ❖ Multiply each digit in the “base” by its place value and add, the check digit will be the number that makes the sum a multiple of eleven. A “digit” of ten is represented by X.

# ISBN Analysis

- Advantages:
  - ❖ This will catch all single digit errors and all transposition errors.
  - ❖ It is not difficult to calculate.
- Disadvantages:
  - ❖ It uses a “strange” character (X) to represent a “digit” (ten).
  - ❖ It works on ten digit numbers.

# ISBN Example

- **0-88385-720-0**

- ❖  $10(0)+9(8)+8(8)+7(3)+6(8)+5(5)+4(7)+3(2)+2(0)+0$

- ❖  $0 + 72 + 64 + 21 + 48 + 25 + 28 + 6$

- ❖  $264 = 11(24)$

- **0-88835-720-0**

- ❖  $72 + 64 + 56 + 18 + 25 + 28 + 6$

- ❖ 269 which is not a multiple of eleven.

# IBM Scheme

- Developed to work with any length identification number.
- Uses a permutation
- Used in credit card numbers.

# IBM Scheme

- Form:

- ❖  $a_1a_2\dots a_{n-1}-c$

- $a_i$  is a digit

- $c$  is the check digit

- The length of the number may be even or odd.

# IBM Scheme Check Digit

- Check:
  - ❖ Apply the permutation in the following manner
  - ❖ For an even length
    - $\sigma(a_1) + a_2 + \dots + \sigma(a_{n-1}) + a_n$
  - ❖ For an odd length
    - $a_1 + \sigma(a_2) + \dots + \sigma(a_{n-1}) + a_n$
  - ❖  $\sigma$  is the permutation **(0)(124875)(36)(9)**
  - ❖ If the number is correct, the sum will be an even multiple of ten.

# IBM Scheme Check Digit

- Calculation:
  - ❖ Apply the permutation to the base number.
  - ❖ If the length of the ID is to be even, apply the permutation to the odd numbered digits, the check digit will be the digit that makes the sum a multiple of ten..
  - ❖ If the length of the ID is to be odd, apply the permutation to the even numbered digits, the check digit will be the digit whose value makes the sum a multiple of ten.

# IBM Scheme Analysis

- Advantages:
  - ❖ May be used with any length number.
  - ❖ Catches all single digit errors.
  - ❖ Catches most transposition errors.
  - ❖ Does not use special characters.
- Disadvantages:
  - ❖ It will not catch transposition errors where nine and zero are transposed.

# IBM Scheme Example

- **4123 4567 8901 2349**

$$\diamond 8 + 1 + 4 + 3 + 8 + 5 + 3 + 7 + 7 + 9 + 0 + 1 + 4 + 3 + 8 + 9 = 80$$

- **5432-0**

$$\diamond 5 + 8 + 3 + 4 + 0 = 20$$

- **4123 4567 8191 2349**

$$\diamond 8 + 1 + 4 + 3 + 8 + 5 + 3 + 7 + 7 + 1 + 9 + 1 + 4 + 3 + 8 + 9 = 81$$

- **5422-0**

$$\diamond 5 + 8 + 2 + 4 + 0 = 19$$

# Verhoeff Scheme

- Goals:
  - ❖ Catch all single digit and transposition errors.
  - ❖ Only use nine digits for the check digit. (No special characters.)
  - ❖ Works for identification numbers of any length.

# Verhoeff Scheme

- Uses the following concepts:
  - ❖ Group
  - ❖ Permutation
  - ❖ Composition of functions
  - ❖ Cayley Table

# Verhoeff Scheme

- Uses the permutation  $\sigma = (0)(14)(23)(56789)$
- Does calculations using the Dihedral Group  $D_{10}$ .
  - ❖ The dihedral group  $D_{10}$  is the collection of symmetries of a regular pentagon.
  - ❖ The binary operation,  $*$ , is the composition of two symmetries.
  - ❖ Each symmetry is represented by a digit.

# Verhoeff Scheme Notation

- The symbol,  $\circ$ , is used to denote composition.
  - ❖  $\sigma \circ \sigma = \sigma^2$
  - ❖  $\sigma \circ \sigma \circ \dots \circ \sigma = \sigma^n$
- The symbol,  $*$ , is used to denote the binary operation in  $D_{10}$ .
  - ❖  $3 * 8 = 6$
  - ❖  $8 * 3 = 5$

# Verhoeff Scheme Cayley Table

*	0	1	2	3	4	5	6	7	8	9
0	0	1	2	3	4	5	6	7	8	9
1	1	2	3	4	0	6	7	8	9	5
2	2	3	4	0	1	7	8	9	5	6
3	3	4	0	1	2	8	9	5	6	7
4	4	0	1	2	3	9	5	6	7	8
5	5	9	8	7	6	0	4	3	2	1
6	6	5	9	8	7	1	0	4	3	2
7	7	6	5	9	8	2	1	0	4	3
8	8	7	6	5	9	3	2	1	0	4
9	9	8	7	6	5	4	3	2	1	0

# Verhoeff Scheme

- Form:
  - ❖  $a_1a_2\dots a_{n-1}-c$ 
    - $a_i$  is a digit
    - $c$  is the check digit.
    - The number may be any length.

# Verhoeff Scheme Check Digit

- Check:
  - ❖  $\sigma^{n-1}(a_1) * \sigma^{n-2}(a_2) * \dots * \sigma(a_{n-1}) * a_n = 0$
- Calculation:
  - ❖ Apply the calculation to the base number, then select the inverse from  $D_{10}$ .

# Verhoeff Scheme Analysis

- Advantages:
  - ❖ Detects all errors listed in the table.
  - ❖ Works with any length number.
  - ❖ Only uses digits (no special characters).
- Disadvantages:
  - ❖ Uses unfamiliar operations.
  - ❖ Not easily calculated manually.