

# Exploring Computer Science

## Data representation

01001010100111101000010010111010010  
001000010100101001001010000101101001010100011101000010010111010010  
11010101010111010000100001010010100100101000010110100101010000111101001010

Adapted from Ms. Martin's "Data" slides by Mr. Bergquist, September 2012  
And from Computer Science Unplugged © 2005 ([www.unplugged.canterbury.ac.nz](http://www.unplugged.canterbury.ac.nz))

# What kind of Files/Data do you keep on...?

- Your Cell Phone?
- A Video Game Console?
- Anything else stored on a Tablet device?
- Garfield School Computers?
  
- What things do you use to save a copy of that data?

# What are the major kinds of information (data) we keep?

- Let's determine these together...

# Computer Data Representation

- How do you think all this Data is represented so a CPU can use it?
  - Hint: It's "digital"...
  - Anyone know what's the alternative to "digital"
    - Music: anyone know how music is stored on vinyl records?
    - Clocks: what's a digital vs an analog clock?

# Binary

- Number system uses two states: on and off
- Any number can be represented!
  - We just need lots of bits (binary digits)
- It is in base 2, but let's review base 10 first and see how it is set up...
- Instead of powers of 10, binary is base 2 uses powers of 2...

# Different representations

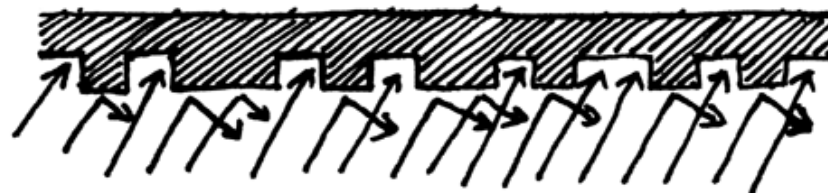
- In RAM, CPU, a bit is power on or off



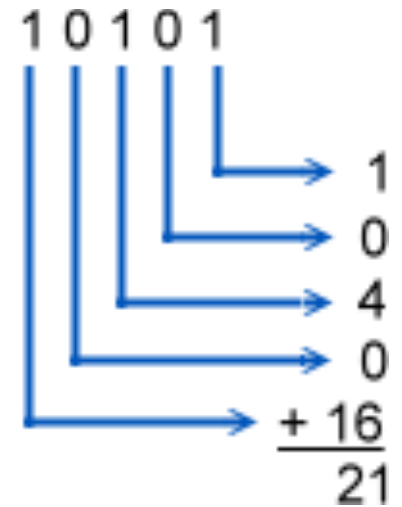
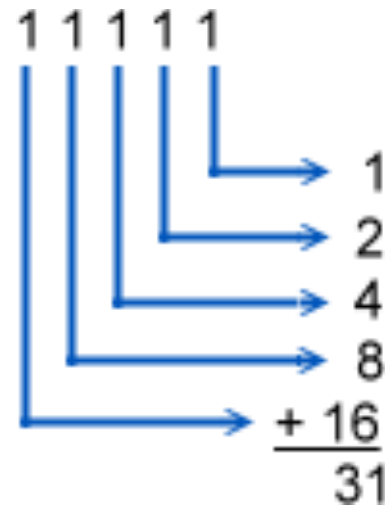
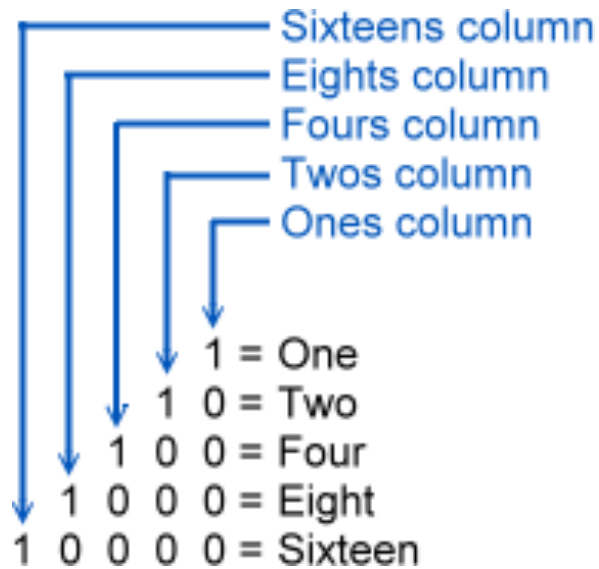
- In hard drive, a bit is magnetic north or south



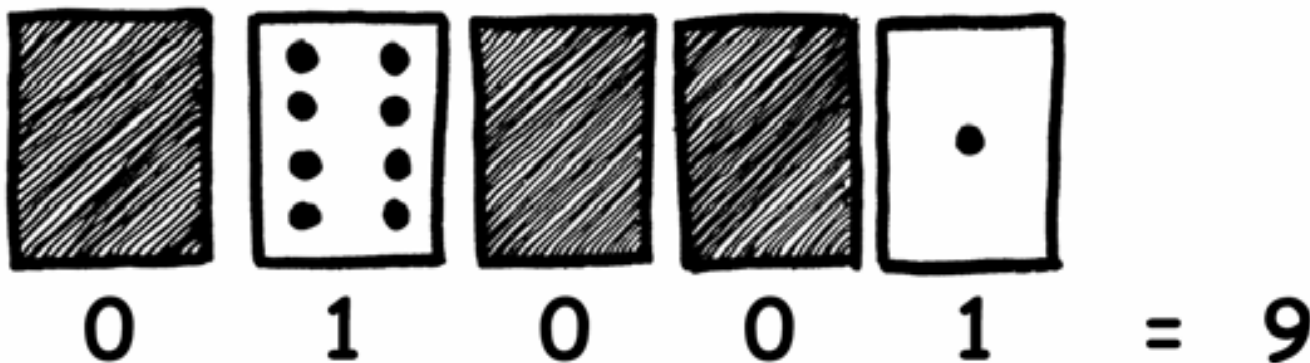
- Audio CDs, CD-ROMs and DVDs store bits optically, a bit either does or does not reflect light



# Base 2 Math...



- But there is an easier way...



# Letters

- Letters are encoded into binary, assigning a number to each
- ASCII - American Standard Code for Information Interchange

## Regular ASCII Chart (character codes 0 - 127)

000	(nul)	016	▶ (dle)	032	sp	048	0	064	@	080	P	096	`	112	p
001	⊙ (soh)	017	◀ (dcl)	033	!	049	1	065	A	081	Q	097	a	113	q
002	⊕ (stx)	018	‡ (dc2)	034	"	050	2	066	B	082	R	098	b	114	r
003	▼ (etx)	019	!! (dc3)	035	#	051	3	067	C	083	S	099	c	115	s
004	⊕ (eot)	020	‡ (dc4)	036	\$	052	4	068	D	084	T	100	d	116	t
005	♣ (enq)	021	§ (nak)	037	%	053	5	069	E	085	U	101	e	117	u
006	⊕ (ack)	022	- (syn)	038	&	054	6	070	F	086	V	102	f	118	v
007	▪ (bel)	023	‡ (etb)	039	'	055	7	071	G	087	W	103	g	119	w
008	■ (bs)	024	† (can)	040	(	056	8	072	H	088	X	104	h	120	x
009	(tab)	025	‡ (em)	041	)	057	9	073	I	089	Y	105	i	121	y
010	(lf)	026	(eof)	042	*	058	:	074	J	090	Z	106	j	122	z
011	♂ (vt)	027	+ (esc)	043	+	059	;	075	K	091	[	107	k	123	{
012	‡ (np)	028	L (fs)	044	,	060	<	076	L	092	\	108	l	124	
013	(cr)	029	↔ (gs)	045	-	061	=	077	M	093	]	109	m	125	}
014	♢ (so)	030	▲ (rs)	046	.	062	>	078	N	094	^	110	n	126	~
015	⊕ (si)	031	▼ (us)	047	/	063	?	079	o	095	_	111	o	127	o



# Letters - extended

- Some names are not that simple, like: Hélène
- Can't be represented in regular ASCII!!
- Extended ASCII, Unicode, and others.

< 妊 > Find  
Aegean\_Numbers

Show Code  
[Help](#)

03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	UTF8
尅	处	夆	夆	备	条	变	女	夂	夂	复	夂	夏	E5A480
5903	5904	5905	5906	5907	5908	5909	590A	590B	590C	590D	590E	590F	
夔	夔	夕	外	夙	夙	夙	多	弓	夜	姓	夔	够	E5A490
5913	5914	5915	5916	5917	5918	5919	591A	591B	591C	591D	591E	591F	
夔	夔	夥	夥	大	夂	天	太	夫	夂	天	夂	夂	E5A4A0
5923	5924	5925	5926	5927	5928	5929	592A	592B	592C	592D	592E	592F	
夂	头	夂	夂	夷	夸	夹	夺	夺	夂	夂	夂	夂	E5A4B0
5933	5934	5935	5936	5937	5938	5939	593A	593B	593C	593D	593E	593F	
夂	奄	夂	夂	奇	奈	奉	夂	夂	卓	夂	夂	夂	E5A580
5943	5944	5945	5946	5947	5948	5949	594A	594B	594C	594D	594E	594F	
夂	奔	夂	奖	套	夂	夂	夂	夂	夂	夂	夂	夂	E5A590
5953	5954	5955	5956	5957	5958	5959	595A	595B	595C	595D	595E	595F	
夂	夂	奥	夂	奥	夂	夂	夂	夂	夂	夂	夂	夂	E5A5A0
5963	5964	5965	5966	5967	5968	5969	596A	596B	596C	596D	596E	596F	
女	奴	夂	奶	奸	奸	她	夂	夂	夂	好	夂	夂	E5A5B0
5973	5974	5975	5976	5977	5978	5979	597A	597B	597C	597D	597E	597F	
妃	妄	夂	妆	妇	妈	夂	妊	夂	夂	妍	夂	夂	E5A680
5983	5984	5985	5986	5987	5988	5989	598A	598B	598C	598D	598E	598F	
妓	夂	夂	妖	夂	夂	妙	夂	夂	夂	妆	夂	夂	E5A690
5993	5994	5995	5996	5997	5998	5999	599A	599B	599C	599D	599E	599F	

# Word document files

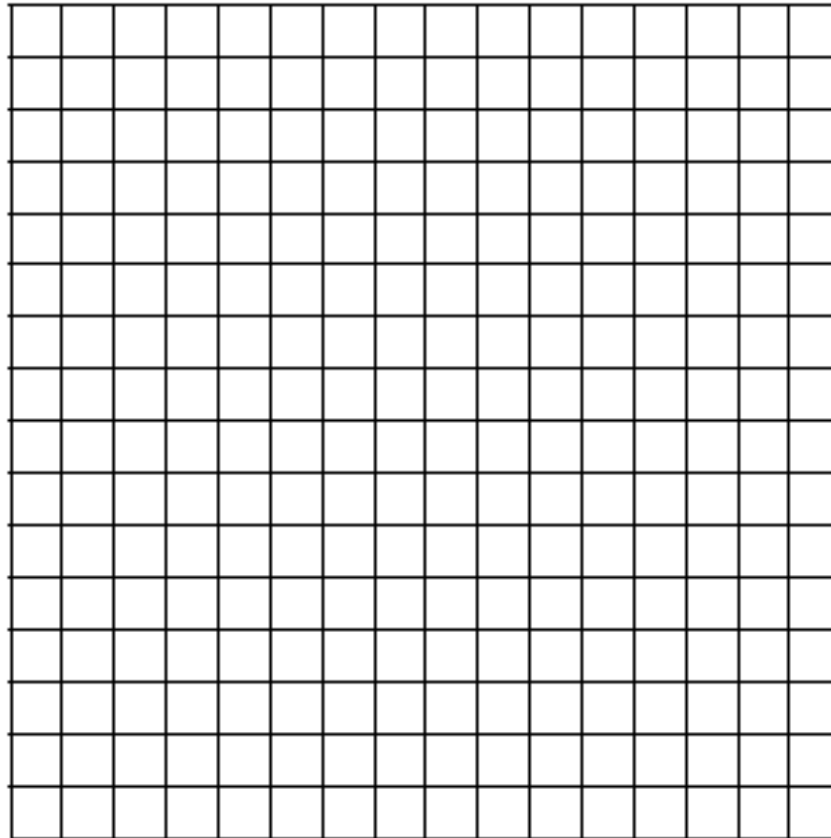
- Store more than just text!
  - There's formatting – when to bold, what font, etc.
  - embedded media
- This "Meta" information takes space
- What else is in our Word Files?

# Knowing how to read characters

- That is the purpose of File extensions like .doc
- Special file headers at the beginning of a file.
- Still sometimes it comes out very wrong!
- Like “Hélène” shows up as:
  - Hilhne - UW student registration
  - HÃ©lÃ¨ne - Zappos
  - Hlllne – other locations, emails

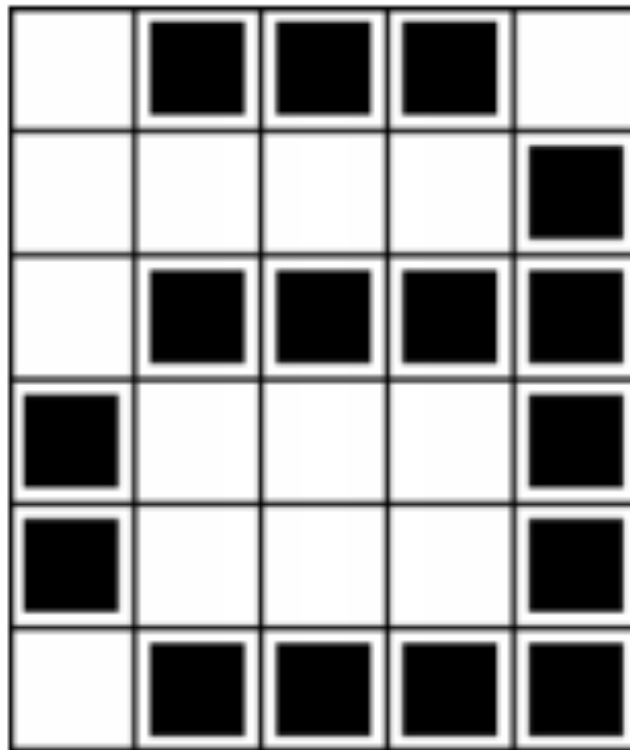
# Images

- All computers know is binary
- Screen is broken down into pixels (picture elements)
- How do we represent images?



# Images: simplified model

- Always start with white count
- Alternate between white and black
- Representing color?



1, 3, 1

4, 1

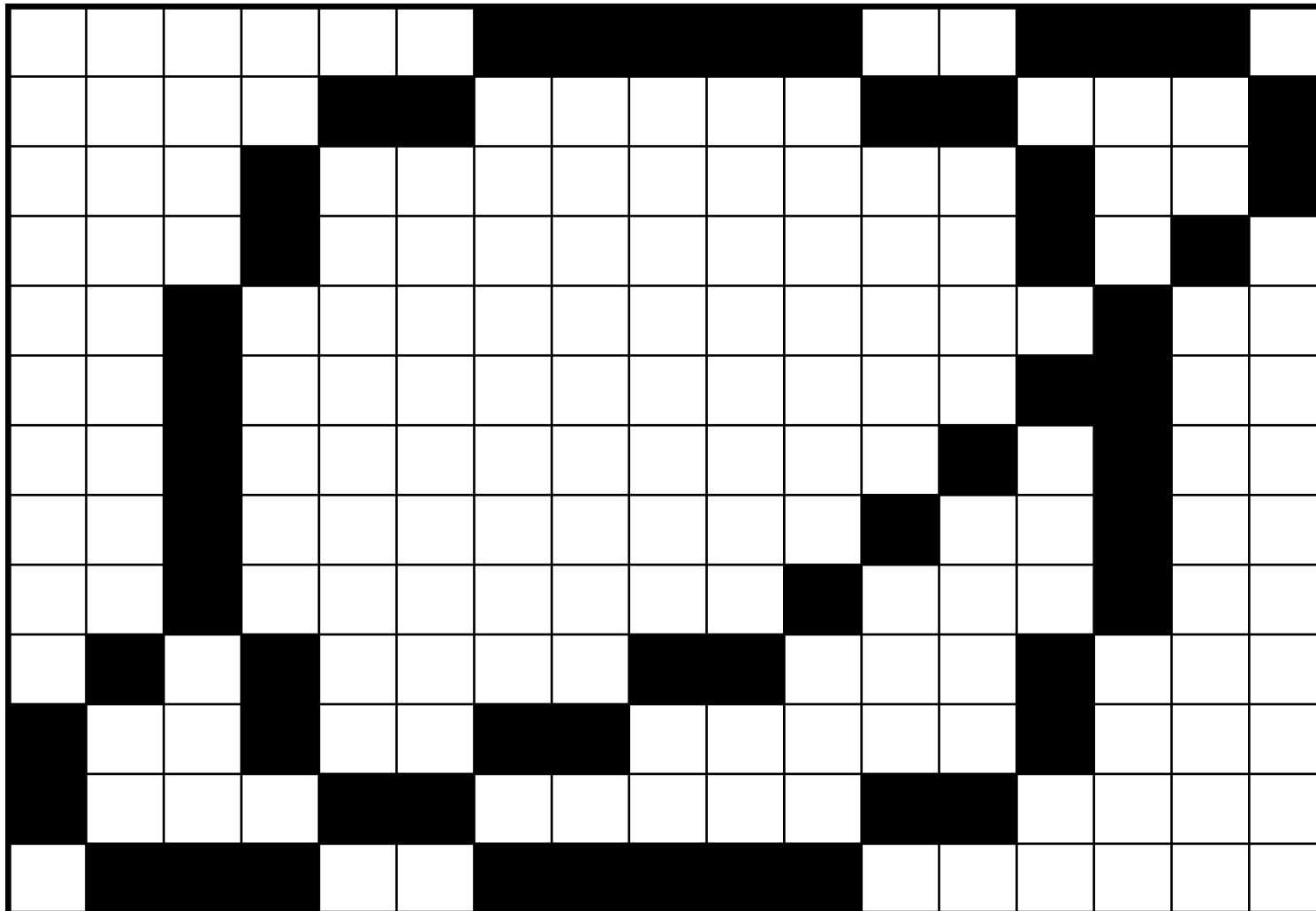
1, 4

0, 1, 3, 1

0, 1, 3, 1

1, 4

# IMAGE BY NUMBERS



6,5,2,3

4,2,5,2,3,1

3,1,9,1,2,1

3,1,9,1,1,1

2,1,11,1

2,1,10,2

2,1,9,1,1,1

2,1,8,1,2,1

2,1,7,1,3,1

1,1,1,1,4,2,3,1

0,1,2,1,2,2,5,1

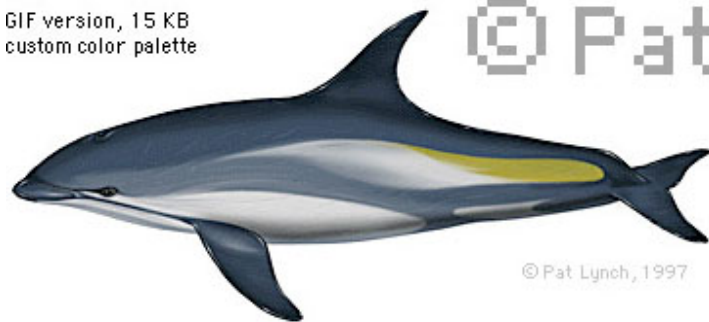
0,1,3,2,5,2

1,3,2,5

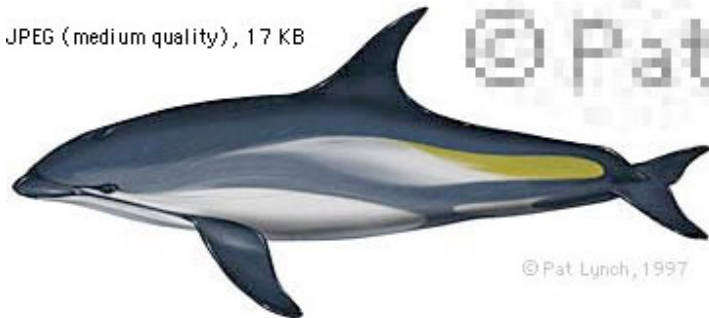
# Image file types

- .bmp or bitmap is close to what we just did
- .jpg, .gif, .png use various forms of compression

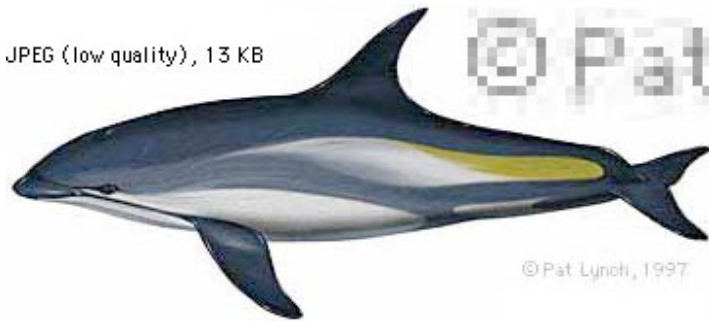
GIF version, 15 KB  
custom color palette



JPEG (medium quality), 17 KB



JPEG (low quality), 13 KB



10100 01000 00001 01110 01011

11001 01111 10101

(THANK YOU)