

# **ГЕНОМ БАКТЕРИИ**

Что можно узнать анализируя  
геном бактерии или археи?

# В СЛОВАРИК

**Прокариоты** – клеточные организмы, не имеющие ядра

Прокариоты делятся на бактерии и археи

**Археи** – вроде бактерий, но эволюционные пути архей и бактерий разошлись 1.5 – 3.5 млрд лет тому назад

**Вирусы** не являются клеточными организмами. Да и организмами их трудно назвать :)

# ИТОГОВОЕ ЗАДАНИЕ ПО БЛОКУ 3

**Дано:** геном бактерии или археи

**Требуется:** Написать мини-обзор о геноме и протеоме этой бактерии или археи

**Методы:**

электронные таблицы, программы пакета EMBOSS, программирование на python

# В СЛОВАРИК

**Геном** – совокупность наследуемой ДНК в клетке организма. Геном многих бактерии состоит из одной кольцевой ДНК. Ее называют **хромосомой**.

У некоторых бактерий бывает две и даже три хромосомы.

У вирусов носителем генома бывает РНК

Часто кроме хромосомы в бактериальной клетке есть еще маленькие ДНК (меньше хромосомы раз в сто), они называются **плазмидами**

# ТАК ВЫГЛЯДИТ ГЕНОМ В ФАЙЛЕ

```
>NC_000964.3 Bacillus subtilis subsp. subtilis str. 168 complete genome
ATCTTTTTCGGCTTTTTTTAGTATCCACAGAGGTTATCGACAACATTTTCACATTACCAACCCCTGTGGA
CAAGGTTTTTTCAACAGGTTGTCCGCTTTGTGGATAAGATTGTGACAACCATTGCAAGCTCTCGTTTTATT
TTGGTATTATATTTGTGTTTTAACTCTTGATTACTAATCCTACSTTTCCTCTTTATCCACAAAGTGTGGA
TAAGTTGTGGATTGATTTACACACAGCTTGTGTAGAAGGTTGTCCACAAGTTGTGAAATTTGTCGAAAAGC
TATTTATCTACTATATTTATATGTTTTCAACATTTAATGTGTACGAATGGTAAGCGCCATTTGCTCTTTTT
TTGTGTTCTATAACAGAGAAAGACGCCATTTTCTAAGAAAAGGAGGGACGTGCCGGAAGATGGAAAATAT
ATTAGACCTGTGGAACCAAGCCCTTGCTCAAATCGAAAAAAGTTGAGCAAACCGAGTTTTTGAGACTTGG
ATGAAGTCAACCAAAGCCCACTCACTGCAAGGCGATACATTAACAATCACGGCTCCCAATGAATTTGCCA
GAGACTGGCTGGAGTCCAGATACTTGCATCTGATTGCAGATACTATATATGAATTAACCGGGGAAGAATT
GAGCATTAAGTTTGTCAATTCCTCAAATCAAGATGTTGAGGACTTTATGCCGAAACCGCAAGTCAAAAAA
GCGGTCAAAGAAGATACATCTGATTTTCTCAAATATGCTCAATCCAAAATATACTTTTGATACTTTTG
TCATCGGATCTGGAAACCGATTTGCACATGCTGCTTCCCTCGCAGTAGCGGAAGCGCCCGCAAAGCTTA
CAACCSTTTATTTATCTATGGGGCGTCGGCTTAGGGAAAACACACTTAATGCATGCGATCGGCCATTAT
GTAATAGATCATAATCSTTCTGCCAAAGTGGTTTATCTGTCTTCTGAGAAATTTACAAACGAATTCATCA
ACTSTATCCGAGATAATAAAGCCGTCGACTTCCGCAATCGSTATCGAAATGTTGATGTGCTTTTGATAGA
TGATATTCAATTTTGTAGCGGGGAAAGAACAACCCAGGAAGAATTTTCCATACATTTAACACATTACAC
GAAGAAAGCAAACAATCGTCATTTCAAGTGACCGGCCGCAAAGGAAATTCGACACTTGAAGACAGAT
```

.....  
.....

После того, как его **секвенировали**, т.е. определили последовательности всех его ДНК

# РАЗМЕРЫ ГЕНОМОВ В ПАРАХ НУКЛЕОТИДОВ (ДНК) ИЛИ БУКВАХ ПО ПОРЯДКУ ВЕЛИЧИН СЛЕДУЕТ ЗНАТЬ

- Вирусы: от – до, пример
- Прокариоты: от – до, пример
- Человек:
- Эукариоты: от – до, пример

Я создал Google Sheet в котором все студенты могут приводить примеры.

Кто найдет самый экстремальный по размеру геном?

# ВОПРОСЫ ПРО ГЕНОМ

# Сегодня вопросы задаете вы

Я записываю вопросы

И подсказываю, каким методом можно  
получить ответ

Вы пробуете найти ответ и приводите его в обзоре

**Какие вопросы можно задать имея  
только текст? (Лингвистика)**



# Лингвистический анализ текста

- Частоты букв
- Часто и редко встречающиеся слова
- Равномерность частоты букв и слов вдоль текста

Эти вопросы изучаются и имеют биологически смысл! Примеры наблюдений:

- $\#C \approx \#G$ ,  $\#T \approx \#A$  ( $\#$  = число)
- Слов  $CG$  *мало* в определенных геномах
- Слов  $TA$  *мало* во всех геномах
- В некоторых геномах  $\#C > \#G$  в одной части и  $\#G > \#C$  в другой части («GC skew»)

# В СЛОВАРИК

**Ген белка** – участок генома, кодирующий один белок. У прокариот почти всегда совпадает с кодирующей последовательностью (**CDS** – Coding Sequence)

**Ген РНК** - участок генома, кодирующий одну молекулу РНК, отличную от матричной РНК

У гена есть

- \* первый нуклеотид
- \* последний нуклеотид
- \* указание цепочки ДНК:
  - + та цепочка, которая лежит в файле с хромосомой;
  - комплементарная цепочка

**Вопрос: направление гена на + и на - цепочках?**

# ВОПРОСЫ ПРО ГЕНЫ

# В СЛОВАРИК

**Протеом:** совокупность белков,  
закодированных в геноме организма

# ТАК ВЫГЛЯДИТ ПРОТЕОМ БАКТЕРИИ В ФАЙЛЕ

**Bacillus cereus strain FORC\_005**, complete genome

GenBank: CP009686.1

LOCUS CP009686 5349617 bp DNA circular BCT 17-NOV-2015

DEFINITION Bacillus cereus strain FORC\_005, complete genome.

ACCESSION **CP009686**

```
gene      1..1341
          /locus_tag="FORC5_0001"
CDS       1..1341
          /locus_tag="FORC5_0001"
          /codon_start=1
          /transl_table=11
          /product="Chromosomal replication initiator protein DnaA"
          /protein_id="AKE14538.1"
          /translation="MENISDLWNSALKELEKKVSKPSYETWLKSTTAHNLKKDVLIT
          .....
          LSRELTDSSLPKIGEEFGGRDHTTVIHAHEKISKLLKTD TQLQKQVEEINGILK"
gene      1520..2665
          /locus_tag="FORC5_0002"
CDS       1520..2665
          /locus_tag="FORC5_0002"
          /codon_start=1
          /transl_table=11
          /product="DNA polymerase III beta subunit"
          /protein_id="AKE14539.1"
          /translation="MRFTIQKDYLVRGVQDVMKAVSSRTTIPILTG IKVVATEEGVTL
          .....
          AKYMMDAL KALDSTEIKVSFTGAMRPFLIRTVNDDSI IQLILPVRTY"
```

# Какие белки бывают у бактерий?

- Нужно импортировать из среды нужные вещества и выкидывать ненужные
  - Это делают белки **транспортеры**, составляющие каналы в мембране. **Трансмембранные** белки
- Нужно реагировать на окружающую среду
  - **Рецепторы** тоже трансмембранные белки, так как передают сигнал из среды в цитоплазму клетки
- Нужно запастись энергией в виде молекул АТФ
  - **АТФсинтазы**, комплексы из многих белков
- Белки, использующие энергию АТФ
  - **АТФазы**

# Какие белки бывают у бактерий?

- Машины для производства белков состоят из белков
  - Машина **РНК полимераза**, состоит из нескольких белков; производит мРНК
  - Машина **рибосома**, состоит из нескольких белков и нескольких РНК
    - Белки называются рибосомальные белки
    - РНК называются рРНК
- Нужно регулировать производство белков

*В Советском Союзе ГОСПЛАН сообщал всем заводам сколько какой продукции надо выпустить в следующем году.*

*У бактерий НЕТ госплана :((((*

- **Транскрипционные факторы** и др.

# Какие белки бывают у бактерий?

- Нужно из субстратов, полученных из среды производить нужные вещества. Например, аминокислоты и нуклеозиды. Нужно модифицировать белки, ДНК, РНК
  - Это делают **ферменты**, их очень много разных
- Нужно размножаться, т.е. делиться на две
  - **ДНК полимераза** и др.
- И еще много-много-много для чего нужны белки
  - вращать жгутики если бактерия умеет передвигаться
  - защищаться от вирусов – бактериофагов (белки иммунитета у бактерий)
  - .....
  - .....



# ВОПРОСЫ ПРО ПРОТЕОМ

# О мини-обзоре

- Структура мини-обзора см. на сайте
- Образцы форматирования см. на сайте

*Зачем навязывают структуры и форматы творческим людям - учёным?*

*Каноны не мешают созданию шедевра*



# ПРОГРАММЫ НА KODOMO

## ПАКЕТ EMBOSS

- seqret
- Infoseq
- wordcount
- fuzznuc
- .....

# ОСНОВЫ EXCEL КАК ИНСТРУМЕНТ ДЛЯ АНАЛИЗА ДАННЫХ

- Из чего состоит ЭТ
- Импорт данных
- Плоская таблица
- Формулы
- vlookup = ВПР

# Дано: Хромосомная таблица – список генов бактерии или археи

```
# feature      class      assembly      assembly_unit seq_type
      chromosome genomic_accession start end strand
      product_accession non-redundant_refseq
      related_accession name symbol GeneID locus_tag
      feature_interval_length product_length
      attributes
gene      protein_coding      GCA_000767275.3      Primary
Assembly      chromosome      CP014334.1      16      1350
      +
      NA23_00005      1335
CDS      with_protein      GCA_000767275.3      Primary Assembly
      chromosome      CP014334.1      16      1350      +
      AMW33602.1      chromosomal replication
initiation protein DnaA      NA23_00005      1335
444
```

# Превратим в Плоскую таблицу

(сейчас, на занятии)

locus_tag	class	feature	genomic_accession	start	end	strand	product_accession	name
EXT02_RS00005	protein_coding	CDS	NZ_CP035949.1	1	1524	+	WP_130427242.1	chromosomal replication initiator protein DnaA
EXT02_RS00010	protein_coding	CDS	NZ_CP035949.1	1819	2955	+	WP_130427244.1	DNA polymerase III subunit beta
EXT02_RS00015	protein_coding	CDS	NZ_CP035949.1	3050	3505	-	WP_130427246.1	hypothetical protein
EXT02_RS00020	protein_coding	CDS	NZ_CP035949.1	4337	4567	+	WP_130427248.1	hypothetical protein
EXT02_RS00025	protein_coding	CDS	NZ_CP035949.1	4698	5063	+	WP_130427250.1	hypothetical protein
EXT02_RS00030	protein_coding	CDS	NZ_CP035949.1	5369	6157	-	WP_130427252.1	hypothetical protein
EXT02_RS00035	protein_coding	CDS	NZ_CP035949.1	7595	8911	+	WP_130427254.1	bifunctional folylpolyglutamate synthase/dihydrofolate synthase
EXT02_RS00040	protein_coding	CDS	NZ_CP035949.1	8982	10190	+	WP_130427256.1	glycosyltransferase family 4 protein
EXT02_RS00045	protein_coding	CDS	NZ_CP035949.1	10363	11610	+	WP_130427258.1	glycosyltransferase family 4 protein
EXT02_RS00050	protein_coding	CDS	NZ_CP035949.1	12158	12445	+	WP_034172155.1	30S ribosomal protein S6
EXT02_RS00055	protein_coding	CDS	NZ_CP035949.1	12460	12804	+	WP_011160385.1	single-stranded DNA-binding protein
EXT02_RS00060	protein_coding	CDS	NZ_CP035949.1	12942	13181	+	WP_011160386.1	30S ribosomal protein S18
EXT02_RS00065	protein_coding	CDS	NZ_CP035949.1	13275	15824	+	WP_130427260.1	50S ribosomal protein L9
EXT02_RS00070	protein_coding	CDS	NZ_CP035949.1	15821	17164	+	WP_130427262.1	replicative DNA helicase

# Сделаем таблицы с результатами (для белков)



Таблица 1. Распределение генов бактерии *Dickeya dadantii* 3937 по категориям.

Тип гена	Число генов
Protein coding	4549
tRNA	75
Pseudogene	22
rRNA	22
ncRNA	13
Antisense RNA	3
RNase P RNA	1
tmRNA	1
SRP RNA	1

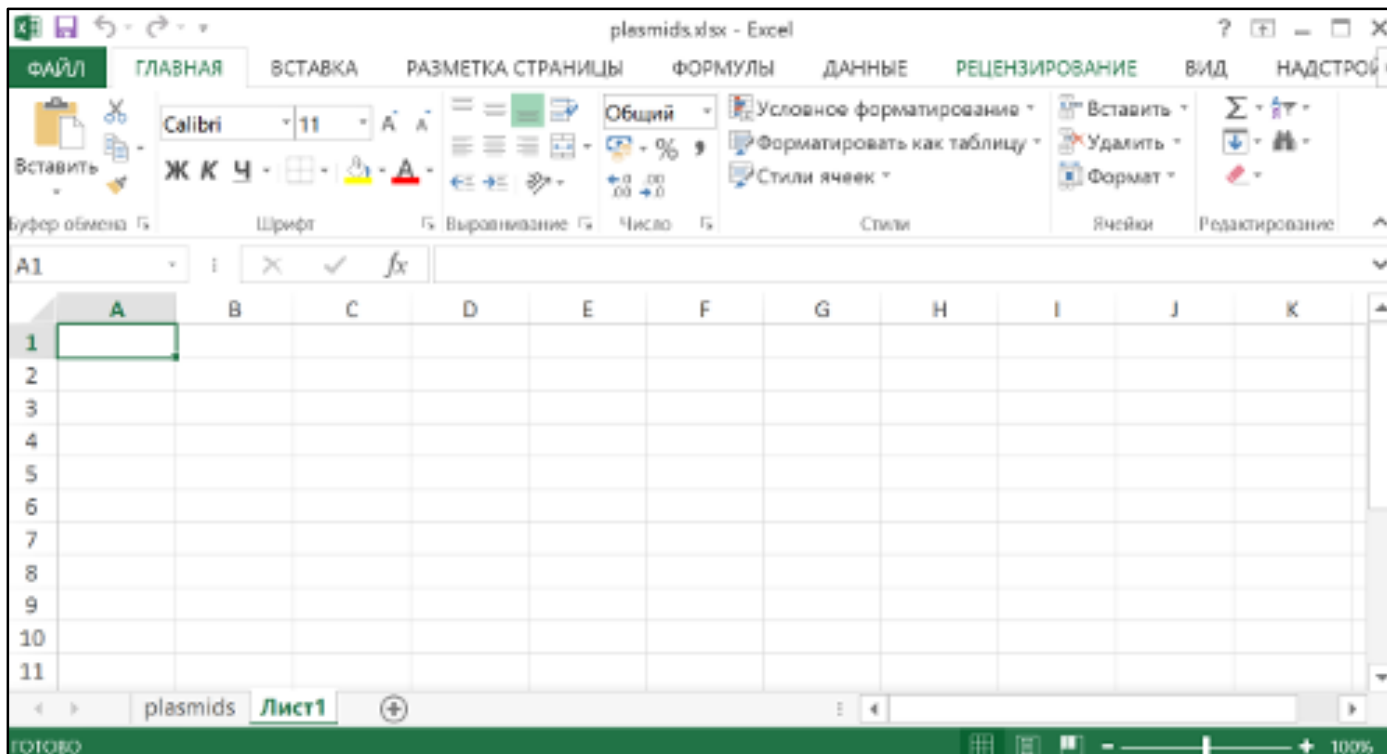


# ИЗ ЧЕГО СОСТОИТ ФАЙЛ EXCEL

Очень бегло – сами разберутся :)

# Файл Excel/Google sheet

- Файл Excel имеет расширение .xlsx или .xls
- Файл состоит из листов
- Лист – таблица; состоит из ячеек



# Ячейка

Адрес:

– **B2**

(в том же листе)

– **TestSheet!B2**

(в том же файле)

– **[TestFile.xlsx]TestSheet!B2**

(в другом файле)

Четыре атрибута  
ячейки:

– значение,

– формула,

– форматирование,

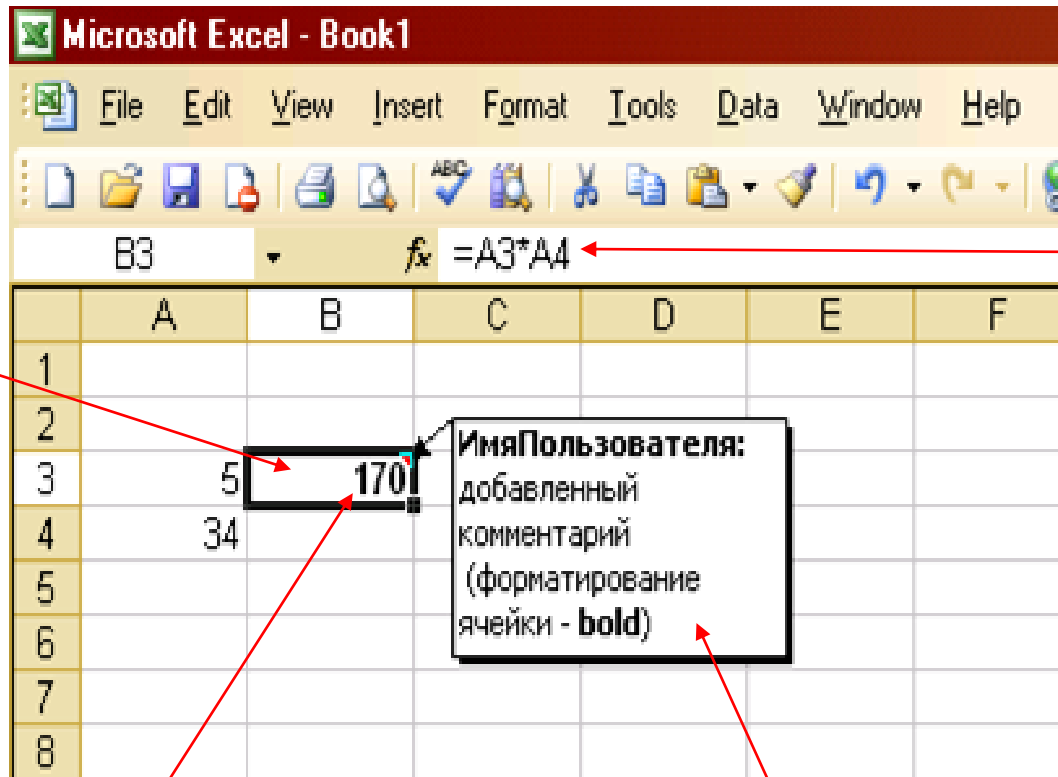
– комментарий

Google Sheet не позволяет  
ссылаться на другие файлы :(

Адрес с \$: **B\$2** или **\$B2** или **\$B\$2**

# Ячейка

Значение:  
170



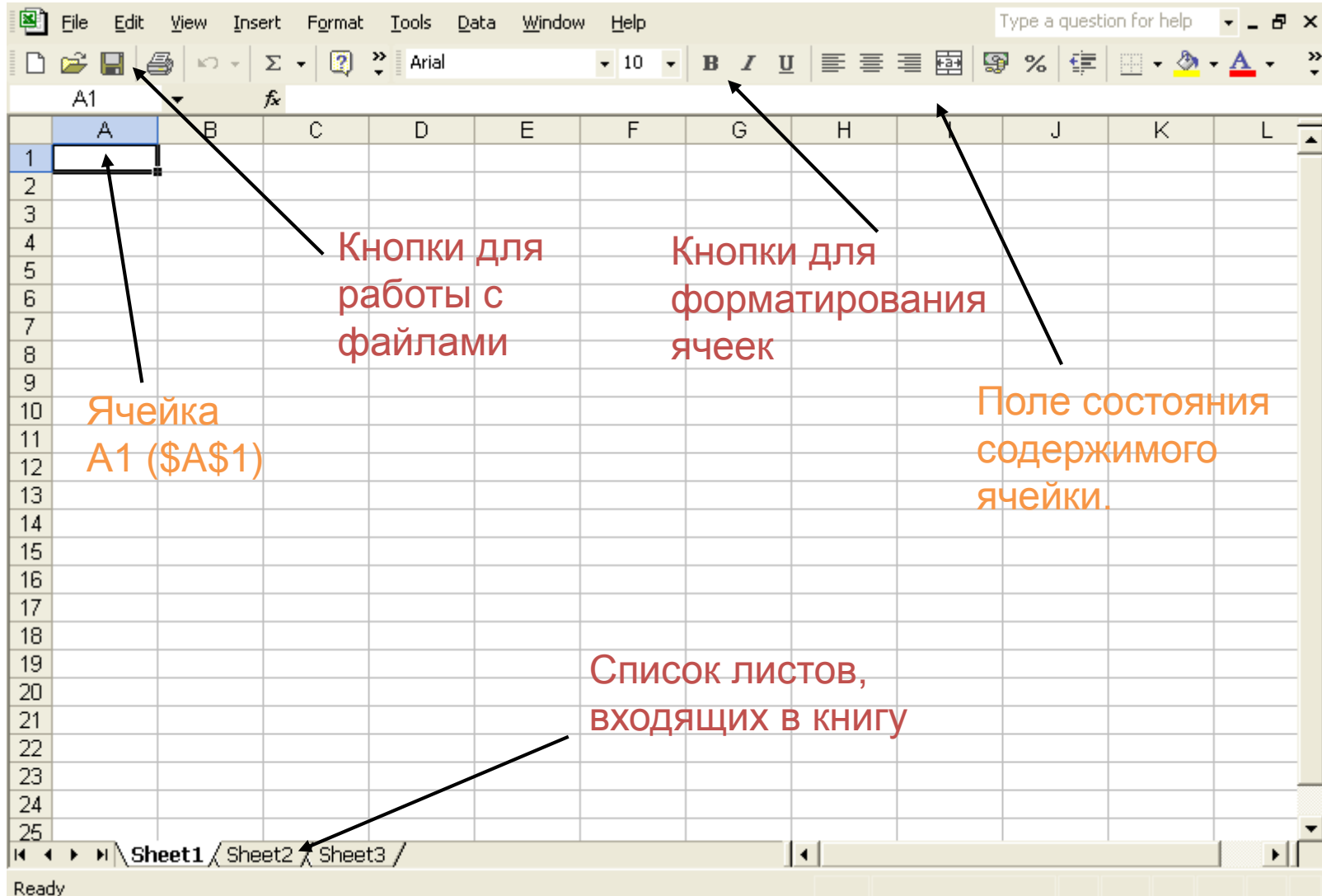
Формула:  
 $=A3*A4$

**Форматирование:**

arial 10, bold,  
правое выравнивание,  
общий формат числа и т.д

**Комментарий**

# Общий вид листа



# Пример в Google Sheet - ведомость

docs.google.com/spreadsheets/d/11goUnRTVgcaqTMjMdaPdXKQMc3FH55HsHy0HQcXHqE/edit#gid=293290325

y20\_results\_term1

File Edit View Insert Format Data Tools Add-ons Help Last edit was 2 hours ago

100% Arial 11

`=semester_results!A1`

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	
	Фамилия	Имя	Группа	Подгруппа	Login	Итог	Зачёт блока	Рейтинг	зачеты прак. и кр.	пр1	пр2	пр3	пр4	пр5	test2.1	КР bash	Рейтинг	пр1	пр2	пр3	пр4	пр5	test2.1	КР bash	1	Зачёт	Рейтинг	1. регистрация	2. wiki страница	3. вопрос	4. письмо	5. координаты гена	Бонус	Штраф	Штраф за опоздание	Когда сдана работа		
2	Weight					да											0.7	1.0	1.0	1.0	1.0	1	1		⇒	2.5	0.7	0.2	0.2	0.2	0.2	1	1	-1	11/09	-		
3	max					7	26.7										21.7	8.0	5.0	5.0	6.5	5.0	5.0	5		4	8.0	5	5	5	5	5	3	3	18/09	-		
4	Алимова	Альфия	2	ПЗ	alfiya.44	ДА	22.7	да	да	да	да	да	да	да	да	да	26.3	3.8	5.0	4.8	6.0	4.4	3.6	3.8		да	3.8	5	5	4	5					09/09		
5	Батин	Артем	2	П4	batin.artem.27	ДА	21.3	да	да	да	да	да	да	да	да	да	25.3	4.0	3.0	4.8	6.0	4.9	4.0	3.2		да	4.0	5	5	5	5	0				08/09		
6	Беляев	Геннадий	1	П2	gennady.belyaev	ДА	24.0	да	да	да	да	да	да	да	да	да	28.0	7.3	5.0	5.0	6.3	3.0	4.0	4.0		з	да	7.3	5	5	4	5	3.5				10/09	
7	Бетенькова	Рената	2	ПЗ	renabet	ДА	25.0	да	да	да	да	да	да	да	да	да	29.6	9.8	5.0	4.8	6.1	2.7	4.7	3.5		н	да	9.8	5	5	4	5	5	1			06/09	
8	Богомаз	Олеся	2	П4	olesjabogomaz	ДА	26.8	да	да	да	да	да	да	да	да	да	31.8	10.0	5.0	4.3	6.0	5.0	5.0	3.6		а	да	10.0	5	5	5	5	5	1			05/09	
9	Бурлака	Артем	1	П1	burlaka.a	ДА	19.3	да	да	да	да	да	да	да	да	да	24.3	4.7	5.0	4.2	4.7	2.3	5.0	4.1		к	да	4.7	5	4.5	2.5	4	4	0.5	2.5	-0.5	11/09	
10	Бурмистрова	Надежда	1	П1	nr.burmistrova		26.6	да	да	да	да	да	да	да	да	да	30.6	9.7	5.0	4.0	6.3	5.0	4.0	3.1		о	да	9.7	5	4.5	4	5	5	1.5			-0.5	14/09
11	Ван	Цзюньчао	2	П4	wangjc																					м	-1.0	5								-2		
12	Васильев	Артем	2	П4	artem.kays	ДА	27.8	да	да	да	да	да	да	да	да	да	32.5	9.8	5.0	5.0	6.5	5.0	4.7	4.1		с	да	9.8	5	5	4	5	5	1			06/09	
13	Винников	Ренат	2	П4	ren-win		21.2	да	да	да	да	да	да	да	да	да	26.2	3.8	5.0	4.5	4.3	5.0	5.0	4.1		т	да	3.8	5	4.5	5	4.5					10/09	
14	Вяльцев	Валерий	1	П1	vyaltsevvaleriy		26.2	да	да	да	да	да	да	да	да	да	30.2	10.5	5.0	4.9	5.4	4.1	4.0	4.2		в	да	10.5	5	4.9	5	5	5	1.5			08/09	
15	Годованец	Ольга	2	ПЗ	o.wgmlx		17.9	да	да		да	да					22.2	8.6	5.0	1.8	1.9	3.6	4.3	3.0		о	да	8.6	5	5	5	5	5	0.1			-0.5	11/09
16	Гуков	Борис	1	П1	jakewayd	ДА	27.4	да	да	да	да	да	да	да	да	да	31.8	9.6	5.0	5.0	6.2	5.0	4.4	3.5		да	9.6	5	4	4	5	5	1				08/09	
17	Давитадзе	Мария	2	П4	mdavitadze		25.9	да	да	да	да	да	да	да	да	да	30.7	9.8	5.0	3.6	6.5	4.4	4.8	4.0		да	9.8	5	4	5	5	5	5	1			10/09	
18	Ефремова	Екатерина	1	П2	e-kat-y.-.		3.3	да	да	да	да	да	да	да	да	да	8.3	2.0	2.0				5.0	2.0		да	2.0	5	5	5	5						-2	01/10
19	Звездин	Дмитрий	2	П4	zvezdindm	ДА	23.1	да	да	да	да	да	да	да	да	да	28.1	8.7	5.0	4.6	4.2	3.6	5.0	3.8		да	8.7	5	4	5	4.5	5						11/09

block1 block2 block3 all\_pracs color-legend preps queue\_service\_sheet tmp

Zayavlenie\_proekt....pdf GCF\_009913275....fna.gz GCF\_000022325....fna.gz

Показать все

# Пример в Google Sheet - ведомость

y20\_results\_term1 ☆ 📁 🌐

File Edit View Insert Format Data Tools Add-ons Help Last edit was 2 hours ago

100% Arial 11 B I U A

=semester\_results!A1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE
1	Фамилия	Имя	Группа	Подгруппа	Login	Угол	Зачёт блока	Рейтинг	зачеты прак. и кр.	пр1	пр2	пр3	пр4	пр5	test2.1	КР bash	Рейтинг	пр1	пр2	пр3	пр4	пр5	test2.1	КР bash	1	Зачёт	Рейтинг	1. регистрация	2. wiki страница	3. вопрос	
2	Weight						да										0.7	1.0	1.0	1.0	1.0	1	1		⇒	2.5	0.7	0.2	0.2	0.2	
3	max						7	26.7									21.7	8.0	5.0	5.0	6.5	5.0	5.0	5		4	8.0	5	5	5	
4	Алимова	Альфия	2	П3	alfiya.44		ДА	22.7		да	да	да	да	да	да	да	26.3	3.8	5.0	4.8	6.0	4.4	3.6	3.8		да	3.8	5	5	4	
5	Батин	Артём	2	П4	batin.artem.27		ДА	21.3		да	да	да	да	да	да	да	25.3	4.0	3.0	4.8	6.0	4.9	4.0	3.2		да	4.0	5	5	5	
6	Беляев	Геннадий	1	П2	gennady.belyaev		ДА	24.0		да	да	да	да	да	да	да	28.0	7.3	5.0	5.0	6.3	3.0	4.0	4.0		з	да	7.3	5	5	4
7	Бетенькова	Рената	2	П3	renabet		ДА	25.0		да	да	да	да	да	да	да	29.6	9.8	5.0	4.8	6.1	2.7	4.7	3.5		н	да	9.8	5	5	4
8	Богомаз	Олеся	2	П4	olesjabogomaz		ДА	26.8		да	да	да	да	да	да	да	31.8	10.0	5.0	4.3	6.0	5.0	5.0	3.6		а	да	10.0	5	5	5
9	Бурлака	Артём	1	П1	burlaka.a		ДА	19.3		да	да	да	да	да	да	да	24.3	4.7	5.0	4.2	4.7	2.3	5.0	4.1		к	да	4.7	5	4.5	2.5
10	Бурмистрова	Надежда	1	П1	nr.burmistrova			26.6		да	да	да	да	да	да	да	30.6	9.7	5.0	4.0	6.3	5.0	4.0	3.1		о	да	9.7	5	4.5	4
11	Ван	Цзюньчао	2	П4	wangjc																					м	-1.0	5			
12	Васильев	Артём	2	П4	artem.kays		ДА	27.8		да	да	да	да	да	да	да	32.5	9.8	5.0	5.0	6.5	5.0	4.7	4.1		с	да	9.8	5	5	4
13	Винников	Ренат	2	П4	ren-win			21.2		да	да	да	да	да	да	да	26.2	3.8	5.0	4.5	4.3	5.0	5.0	4.1		т	да	3.8	5	4.5	5
14	Вяльцев	Валерий	1	П1	vyaltsevvaleriy			26.2		да	да	да	да	да	да	да	30.2	10.5	5.0	4.9	5.4	4.1	4.0	4.2		в	да	10.5	5	4.9	5
15	Годованец	Ольга	2	П3	o.wgmlx			17.9		да	да	да	да	да	да	да	22.2	8.6	5.0	1.8	1.9	3.6	4.3	3.0		о	да	8.6	5	5	5
16	Гуков	Борис	1	П1	jakewayd		ДА	27.4		да	да	да	да	да	да	да	31.8	9.6	5.0	5.0	6.2	5.0	4.4	3.5		.	да	9.6	5	4	4
17	Давитадзе	Мария	2	П4	mdavitadze			25.9		да	да	да	да	да	да	да	30.7	9.8	5.0	3.6	6.5	4.4	4.8	4.0		да	9.8	5	4	5	
18	Ефремова	Екатерина	1	П2	e-kat-y.-.			3.3		да	да	да	да	да	да	да	8.3	2.0	2.0				5.0	2.0		да	2.0	5	5	5	
19	Звездин	Дмитрий	2	П4	zvezdindm		ДА	23.1		да	да	да	да	да	да	да	28.1	8.7	5.0	4.6	4.2	3.6	5.0	3.8		да	8.7	5	4	5	

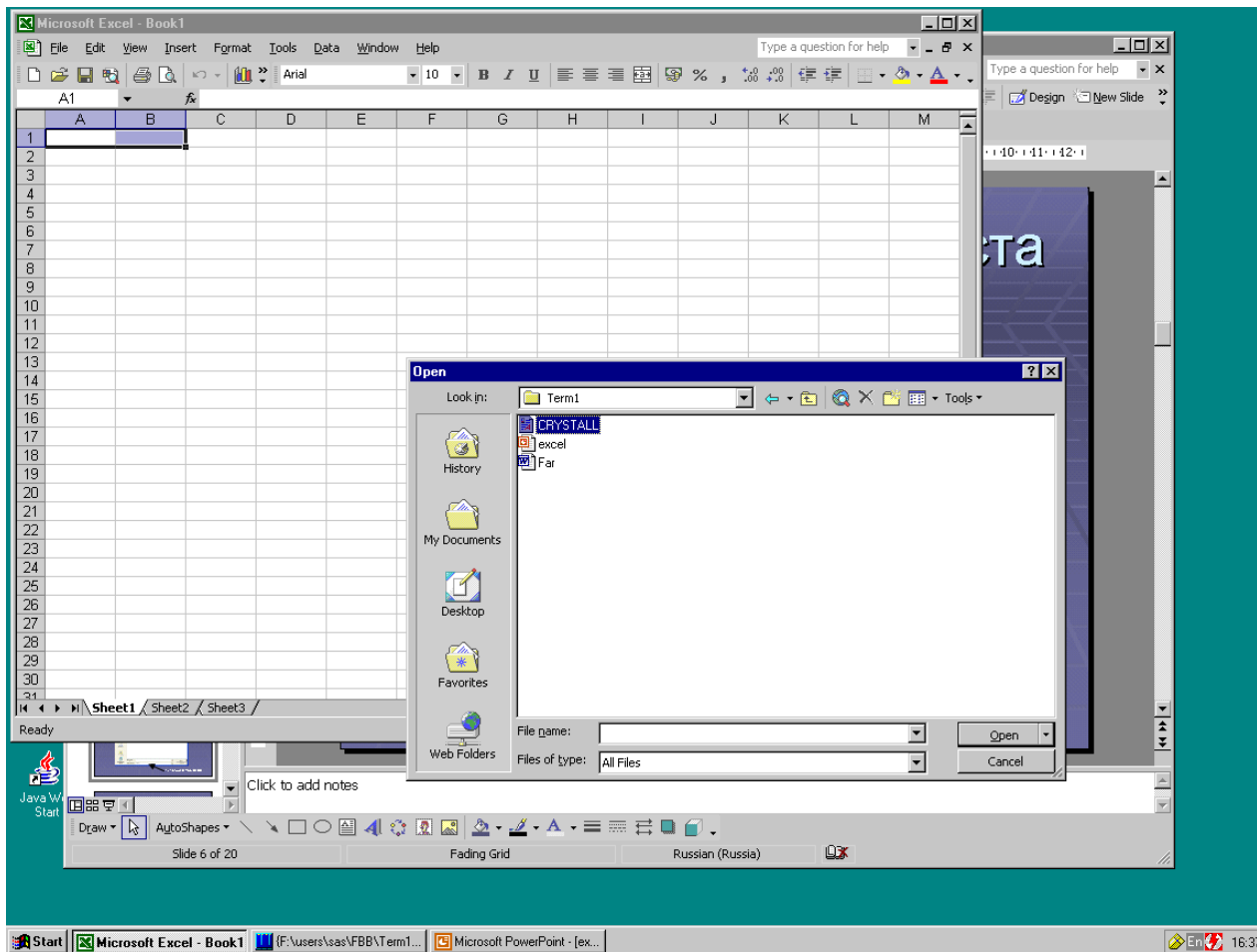
+ \_results queue Q&A block1 block2 block3 all\_pracs color-legend preps queue\_service\_sheet tmp

Zayavlenie\_proekt....pdf Zayavlenie\_proekt....pdf GCF\_009913275....fna.gz GCF\_000022325....fna.gz

# ИМПОРТ ДАННЫХ



# Импорт данных из текста



# Импорт данных из текста

Тип текстовой таблицы

Кодировка

**Text Import Wizard - Step 1 of 3**

The Text Wizard has determined that your data is Fixed Width.  
If this is correct, choose Next, or choose the data type that best describes your data.

Original data type  
Choose the file type that best describes your data:

- Delimited - Characters such as commas or tabs separate each field.
- Fixed width - Fields are aligned in columns with spaces between each field.

Start import at row: 25 File origin: 1254 : Turkish (Windows)

Preview of file F:\users\sas\FBB\Term1\CRYSTALL.pdb.

23	SCALE2	0.000000	0.021316	0.000000	0.000000						
24	SCALE3	0.000000	0.000000	0.006089	0.000000						
25	ATOM	1	N	LYS	A	74	18.732	8.253	-20.201	1.00	9
26	ATOM	2	CA	LYS	A	74	18.339	9.325	-19.243	1.00	9
27	ATOM	3	C	LYS	A	74	18.492	8.867	-17.798	1.00	9

Cancel < Back Next > Finish

# Импорт данных из текста

Разделители данных

Text Import Wizard - Step 2 of 3

This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below.

Delimiters

Tab  Semicolon  Comma

Space  Other:

Treat consecutive delimiters as one

Text qualifier: "  "

Data preview

SCALE2	0.000000	0.021316	0.000000	0.000000				
SCALE3	0.000000	0.000000	0.006089	0.000000				
ATOM	1	N	LYS	A	74	18.732	8.253	-20.2
ATOM	2	CA	LYS	A	74	18.339	9.325	-19.2
ATOM	3	C	LYS	A	74	18.492	8.867	-17.7

Cancel < Back Next > Finish

# Импорт данных из текста

Microsoft Excel - Book1

Type a question for help

Arial 10 B I U

A1

Text Import Wizard - Step 3 of 3

This screen lets you select each column and set the Data Format.

'General' converts numeric values to numbers, date values to dates, and all remaining values to text.

Advanced...

Column data format

- General
- Text
- Date: DMY
- Do not import column (skip)

Data preview

	Gener	Cener	Cener	Cener	Cener	Gener	Gener	General	Gener	Gener	Gener
ATOM 1	1	N	LYS	A	74	18.732	8.253	-20.201	1.00	92.76	N
ATOM 2	2	CA	LYS	A	74	18.339	9.325	-19.243	1.00	92.81	C
ATOM 3	3	C	LYS	A	74	18.492	8.867	-17.798	1.00	93.32	C
ATOM 4	4	D	LYS	A	74	18.009	7.798	-17.422	1.00	93.66	D
ATOM 5	5	CB	LYS	A	74	16.904	9.750	-19.500	1.00	90.24	C

Cancel < Back Next > Finish

Sheet1 Sheet2 Sheet3

Ready NUM SCRL

Определение  
типа данных  
в колонке

# Импорт данных из текста

Microsoft Excel - Crystal

File Edit View Insert Format Tools Data Window Help

Type a question for help

K802 29.06.2004

	A	B	C	D	E	F	G	H	I	J	K	L	M
781	ATOM	781	C	LYS	B	177	-11.358	14.982	2.784	1.00	35.38	C	
782	ATOM	782	O	LYS	B	177	-10.387	15.202	3.509	1.00	33.89	O	
783	ATOM	783	CB	LYS	B	177	-13.415	16.431	2.836	1.00	36.37	C	
784	ATOM	784	CG	LYS	B	177	-13.993	17.791	2.461	1.00	37.76	C	
785	ATOM	785	CD	LYS	B	177	-15.468	17.927	2.812	1.00	42.33	C	
786	ATOM	786	CE	LYS	B	177	-15.717	17.914	4.309	1.00	44.59	C	
787	ATOM	787	NZ	LYS	B	177	-17.165	18.102	4.605	1.00	47.49	N	
788	ATOM	788	N	ASN	B	178	-11.807	13.757	2.533	1.00	36.51	N	
789	ATOM	789	CA	ASN	B	178	-11.162	12.587	3.110	1.00	36.13	C	
790	ATOM	790	C	ASN	B	178	-9.721	12.485	2.633	1.00	34.95	C	
791	ATOM	791	O	ASN	B	178	-8.830	12.129	3.402	1.00	32.72	O	
792	ATOM	792	CB	ASN	B	178	-11.934	11.319	2.739	1.00	38.55	C	
793	ATOM	793	CG	ASN	B	178	-12.971	10.952	3.779	1.00	40.52	C	
794	ATOM	794	OD1	ASN	B	178	-14.111	10.623	3.453	1.00	43.43	O	
795	ATOM	795	ND2	ASN	B	178	-12.573	10.999	5.044	1.00	43.93	N	
796	ATOM	796	N	TRP	B	179	-9.487	12.809	1.366	1.00	32.96	N	
797	ATOM	797	CA	TRP	B	179	-8.134	12.741	0.839	1.00	30.78	C	
798	ATOM	798	C	TRP	B	179	-7.227	13.733	1.567	1.00	33.00	C	
799	ATOM	799	O	TRP	B	179	-6.110	13.390	1.949	1.00	35.23	O	
800	ATOM	800	CB	TRP	B	179	-8.113	13.040	-0.663	1.00	29.30	C	
801	ATOM	801	CG	TRP	B	179	-6.775	12.755	-1.280	1.00	29.83	C	
802	ATOM	802	CD1	TRP	B	179	-6.337	11.560	-1.773	1.00	29.июн	C	
803	ATOM	803	CD2	TRP	B	179	-5.672	13.665	-1.390	1.00	27.85	C	
804	ATOM	804	NE1	TRP	B	179	-5.028	11.667	-2.182	1.00	28.59	N	
805	ATOM	805	CE2	TRP	B	179	-4.599	12.947	-1.962	1.00	26.55	C	
806	ATOM	806	CE3	TRP	B	179	-5.492	15.014	-1.066	1.00	27.59	C	
807	ATOM	807	CZ2	TRP	B	179	-3.351	13.536	-2.205	1.00	27.83	C	
808	ATOM	808	CZ3	TRP	B	179	-4.252	15.600	-1.310	1.00	29.64	C	
809	ATOM	809	CH2	TRP	B	179	-3.200	14.859	-1.877	1.00	27.58	C	
810	ATOM	810	N	VAL	B	180	-7.709	14.956	1.774	1.00	32.28	N	
811	ATOM	811	CA	VAL	B	180	-6.908	15.980	2.442	1.00	30.59	C	

CRYSTALL

Ready NUM SCRL

# ПЛОСКАЯ ТАБЛИЦА

# Какая таблица лучше для технологии работы с ней?

ПЕРИОДЫ	ГРУППЫ ЭЛЕМЕНТОВ	ГРУППЫ ЭЛЕМЕНТОВ																
		A I B	A II B	A III B	A IV B	A V B	A VI B	A VII B	A VIII B									
I	1	H	ПЕРИОДИЧЕСКАЯ СИСТЕМА ЭЛЕМЕНТОВ Д. И. МЕНДЕЛЕЕВА (H)						He									
II	2	Li	Be	B	C	N	O	F	Ne									
III	3	Na	Mg	Al	Si	P	S	Cl	Ar									
IV	4	K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni							
V	5	Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd							
VI	6	Cs	Ba	La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
VII	7	Fr	Ra	Ac	Rf	Db	Sg	Bh	Hs	Mt	Ds							
ВЫСШИЕ ОКСИДЫ		RO	RO	RO <sub>2</sub>	RO <sub>2</sub>	RO <sub>2</sub>	RO <sub>3</sub>	RO <sub>3</sub>	RO <sub>3</sub>	RO <sub>4</sub>	RO <sub>4</sub>							
ЛЕТУЧЕ ВОДОРОДНЫЕ СОЕДИНЕНИЯ				RH <sub>4</sub>	RH <sub>4</sub>	RH <sub>3</sub>	RH <sub>3</sub>	RH <sub>2</sub>	RH <sub>2</sub>	RH	RH							
* ЛАНТАНОИДЫ																		
* АКТИНОИДЫ																		

	A	B	C	D	E
1	Результаты контрольной дойки				
2	Кличка коровы	Надой, кг	МДЖ, %	МДБ, %	
3	Акация	18,3	3,45	3,6	
4	Белуга	25,2	3,1	3,2	
5	Волна	16,3	3,12	3,22	
6	Гамма	15,9	3,26	3,38	
7	Легенда	14,9	3,64	3,53	
8	Линейка	26,1	3,45	3,2	
9	Перепелица	12,9	3,78	3,57	
10	Свободная	11,3	4,01	3,9	
11	Сильная	18,9	3,71	3,62	
12	Сирень	16,3	3,12	3,41	
13	Итого	176,1	3,43	3,42	
14					
15					

# Плоская таблица

Genome	Super-kingdom	no rank	phylum	class
CP001787.1	Archaea		Euryarchaeota	Methanococci
CP002737.1	Archaea	Terrabacteria group	Euryarchaeota	Methanococci
L77117.1	Archaea		Euryarchaeota	Methanococci
CP013050.1	Archaea		Euryarchaeota	Thermococci
BA000011.4	Archaea		Euryarchaeota	Thermoplasmata
DP000238.1	Archaea	TACK group	Thaumarchaeota	
CP003081.1	Bacteria	Bacteroidetes/Chlorobi group	Actinobacteria	Actinobacteria
CP011499.1	Bacteria	Terrabacteria group	Actinobacteria	Actinobacteria
CP001080.1	Bacteria		Aquificae	Aquificae



# Требования

- Названия колонок – в первой строке. Если названия длинные, то удобно повернуть текст вертикально и разрешить перенос.
- Остальные строки – с данными.
- Все ячейки заполнены! Пустая ячейка может быть только если это – предусмотренное значение; например, обозначает «данные неизвестны»
- В первой колонке – уникальный идентификатор строки
- **НЕТ ОБЪЕДИНЕНИЙ ЯЧЕЕК!**

# Не плоская таблица. Пример.

Как надо было ее составить?

Microsoft Excel - Книга1

Файл Правка Вид Вставка Формат Сервис Данные Окно ?

	A	B	C	D
1	<b>Перечень техники, приобретенной кафедрами энергетического факультета ИрГТУ в 2000 г.</b>			
2	<b>Наименование</b>	<b>Кол-во</b>	<b>Цена</b>	<b>Сумма</b>
3	Кафедра теплоэнергетики			
4	Компьютеры	11	17800	195800
5	Сканеры	1	3400	3400
6	Звуковые колонки	2	200	400
7	<b>Итого:</b>			<b>199600</b>
8	Кафедра электрических станций, сетей и систем			
9	Компьютеры	10	17800	178000
10	Принтеры	1	5700	5700
11	Сканеры	1	3400	3400
12	<b>Итого:</b>			<b>187100</b>
13	Кафедра электропривода и автоматизации			
14	Компьютеры	4	18700	74800
15	Принтеры	1	6900	6900
16	<b>Итого:</b>			<b>81700</b>
17				
18	<b>Всего:</b>			<b>468400</b>
19				

Исходная таблица Лист2 Лист3

Готово NUM

Показать на примере хромосомной таблицы

# **ВПР (= VLOOKUP) – ВЕРТИКАЛЬНЫЙ ПРОСМОТР**

# ВПР (VLOOKUP)

Microsoft Excel - Book1

File Edit View Insert Format Tools Data Window Help

D4 fx

	A	B	C
1	Аминокислота, иминокислота пролин	Масса amino-(имино-) кислоты	
2	A	89,09	
3	C	121,16	
4	D	133,1	
5	E	147,13	
6	F	165,19	
7	G	75,07	
8	H	155,1	
9	I	131,17	
10	K	146,19	
11	L	131,17	
12	M	149,21	
13	N	132,12	
14	P	115,13	
15	Q	146,15	
16	R	174,2	
17	S	105,09	
18	T	119,12	
19	V	117,15	
20	W	204,22	
21	Y	181,19	
22			
23			
24			
25			

1

Microsoft Excel - Book1

File Edit View Insert Format Tools Data Window Help

C3 fx

	A	B	C	D
1				
2		А/к последовательность	молекулярная масса а/к остатка	
3		L		
4		V		
5		A		
6		L		
7		V		
8		R		
9		A		
10		S		
11		I		
12		T		
13		D		
14		P		
15		G		
16		R		
17		L		
18		P		
19		E		
20		N		
21		P		
22		K		
23		I		
24		P		
25		H		
26		G		
27		E		
28				

2

# ВПР (VLOOKUP)

## Искомое значение

Искомое значение (lookup\_value)

адрес интервала ячеек,  
содержащих идентификатор (в  
данном случае, код остатка),  
который нужно найти в первой  
колонке 1й таблицы

\*Ссылка указывает название листа и  
интервал данных

The screenshot shows an Excel spreadsheet with a VLOOKUP formula in cell C3. The formula is `=VLOOKUP(Sheet1!B3:B27)`. The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I
1									
2		А/к последова тельность	молекулярная масса а/к остатка						
3		L	=VLOOKUP(Sheet1!B3:B27)						
4		V							
5		A							
6		L							
7		V							
8		R							
9		A							
10		S							
11		I							
12		T							
13		D							
14		P							
15		G							
16		R							
17		L							
18		P							
19		E							
20		N							
21		P							
22		K							
23		I							
24		P							
25		H							
26		G							
27		E							

The Function Arguments dialog box for VLOOKUP is open, showing the following fields:

- Lookup\_value: Sheet1!B3:B27
- Table\_array: (empty)
- Col\_index\_num: (empty)
- Range\_lookup: (empty)

The dialog box also includes a description of the function and a 'Formula result =' field.

# ВПР (VLOOKUP) таблица

## таблица (table\_array)

адрес таблицы (1), в первой колонке которой следует искать заданный идентификатор (а/к)

номер столбца (Col\_index\_num)  
из табл.1, который содержит необходимые значения для табл.2 (второй столбец)

интервальный просмотр (range lookup)  
указать ЛОЖЬ, чтобы всё получилось ;)

The screenshot shows an Excel spreadsheet with a table of amino acids and their masses. The formula bar displays `=VLOOKUP(Sheet1!B3:B27;aa_mass!A2:B21;2;FALSE)`. The table has two columns: 'Аминокислота, иминокислота пролин' and 'Масса аминок-(имино-) кислоты'. The 'Function Arguments' dialog box is open, showing the following values: Lookup\_value: Sheet1!B3:B27, Table\_array: aa\_mass!A2:B21, Col\_index\_num: 2, Range\_lookup: FALSE. The result of the formula is shown as 131,17.

	A	B	C	D	E	F	G
1	Аминокислота, иминокислота пролин	Масса аминок-(имино-) кислоты					
2	A	89,09					
3	C	121,16					
4	D	133,1					
5	E	147,13					
6	F	165,19					
7	G	75,07					
8	H	155,1					
9	I	131,17					
10	K	146,19					
11	L	131,17					
12	M	149,21					
13	N	132,12					
14	P	115,					
15	Q	146,					
16	R	174,					
17	S	105,					
18	T	119,					
19	V	117,					
20	W	204,					
21	Y	181,					

**Function Arguments**

VLOOKUP

Lookup\_value: Sheet1!B3:B27 = "L"

Table\_array: aa\_mass!A2:B21 = {"A";89,09;"C";121,16;"D";133,1;"E";147,13;"F";165,19;"G";75,07;"H";155,1;"I";131,17;"K";146,19;"L";131,17;"M";149,21;"N";132,12;"P";115,;"Q";146,;"R";174,;"S";105,;"T";119,;"V";117,;"W";204,;"Y";181,}

Col\_index\_num: 2 = 2

Range\_lookup: FALSE = FALSE

Formula result = 131,17

Looks for a value in the leftmost column of a table, and then returns a value in the same row from a column you specify. By default, the table must be sorted in an ascending order.

Table\_array is a table of text, numbers, or logical values, in which data is retrieved. Table\_array can be a reference to a range or a range name.

OK Cancel

# ВПР (VLOOKUP)

результат

распространение

fx =VLOOKUP(Sheet1!B3:B27;aa\_mass!A2:B21;2;FALSE)

A/к последовательность	молекулярная масса а/к остатка
L	131,17
V	
A	
L	
V	
R	
A	
S	
I	
T	
D	
P	
G	
R	
L	
P	
E	
N	
P	
K	
I	
P	
H	
G	
E	

fx =VLOOKUP(Sheet1!\$B\$3:\$B\$27;aa\_mass!\$A\$2:\$B\$21;2;FALSE)

A/к последовательность	молекулярная масса а/к остатка
L	131,17
V	117,15
A	89,09
L	131,17
V	117,15
R	174,2
A	89,09
S	
I	
T	
D	
P	
G	
R	
L	
P	
E	
N	
P	
K	
I	
P	
H	
G	
E	

\$ закрепить области!!!!

**КОНЕЦ**

Пока остановимся ...



# **СПРАВОЧНЫЙ МАТЕРИАЛ**

# Формула

- Ввод – через “=”
- Встроенные функции
- Мастер формул

	A	B	C	D	E
1					
2	A0007	303	100		
3	A0008	261			
4	A0009	341			
5	A0010	600			

уфер обмена | Шрифт | Категория: Математические

D3 | fx | =

Выберите функцию:

- СУММРАЗМКВ
- СУММСУММКВ
- ФАКТР
- ЦЕЛОЕ
- ЧАСТНОЕ**
- ЧЁТН
- ЧИСЛКОМБ

**ЧАСТНОЕ(числитель,знамен)**  
Возвращает целую часть резу

	A	B	C	D
1				
2	A0007	303	100	
3	A0008	261	86	=
4	A0009	341	112.67	112
5	A0010	600	199	

	A	B	C	D	E	F
1						
2	A0007	303	100			
3	A0008	261	86			
4	A0009	341	112.67	112		
5	A0010	600	199			

# Встроенные функции

## Математические

- **abs**
- **макс**
- **sin**
- **мин**
- **exp**
- **целое**
- **остат**
- **частное**
- **Log**
- **корень**

## Текстовые

- **сжпробелы**
- **левсимв**
- **правсимв**
- **пстр**
- **сцепить**
- **считатьпустоты**
- **длстр**
- **найти**
- **поиск**

## Статистические

- **срзнач**
- **стандотклон**
- **медиана**
- **случмежду**

## Ссылки

- **впр**
- **гпр**
- **строка**
- **столбец**
- **индекс**

## Проверка значений

- **счётесли**
- **счётеслимн**
- **суммесли**
- **суммеслимн**

## Логические

- **если**
- **И**
- **ИЛИ**
- **НЕ**
- **еслиошибка**
- **ИСТИНА**
- **ЛОЖЬ**

... и много других

# Распространение формулы

- Задача: вычислить интеграл от 0 до 1 от  $e^x$

A3					$=A2+E\$1$
	A	B	C	D	E
1	x	$e^x$	интеграл	step =	0.01
2	0	1			
3	0.01				
4					

**Ctrl+D**

Доллар запрещает изменение при распространении

A4					$=A3+E\$1$	
	A	B	C	D	E	F
1	x	$e^x$	интеграл	step =	0.01	
2	0	1				
3	0.01					
4	0.02					
5						

A3:A101					$=A2+E\$1$	
	A	B	C	D	E	F
3	0.01					
4						
5						
6						
7						

**Ctrl+D**

Выделение прямоугольной области

Распространение вниз

(Ctrl+R – распространение вправо)

# ... продолжение

B2				fx		=EXP(A2)	
	A	B	C	D	E	F	
1	x	e^x	интеграл	step =	0.01		
2	0	1					
3	0.01						
4	0.02						
5	0.03						

Формула для  
интеграла от 0  
до x

Двойной щелчок приводит  
к распространению до  
последней заполненной  
ячейки в соседнем столбце

C5				fx		=СУММ(B\$2:B5)*E\$1	
	A	B	C	D	E	F	
1	x	e^x	интеграл (0,x)	step =	0.01		
2	0	1	0.01				
3	0.01	1.0101	0.020100502				
4	0.02	1.0202	0.030302515				
5	0.03	1.0305	0.04060706				
6	0.04	1.0408	0.051015168				

# Фиксировать можно и столбцы

- Знак \$:

**K\$29:L30** или **K\$29:\$L30** или **\$K\$29:\$L\$30**

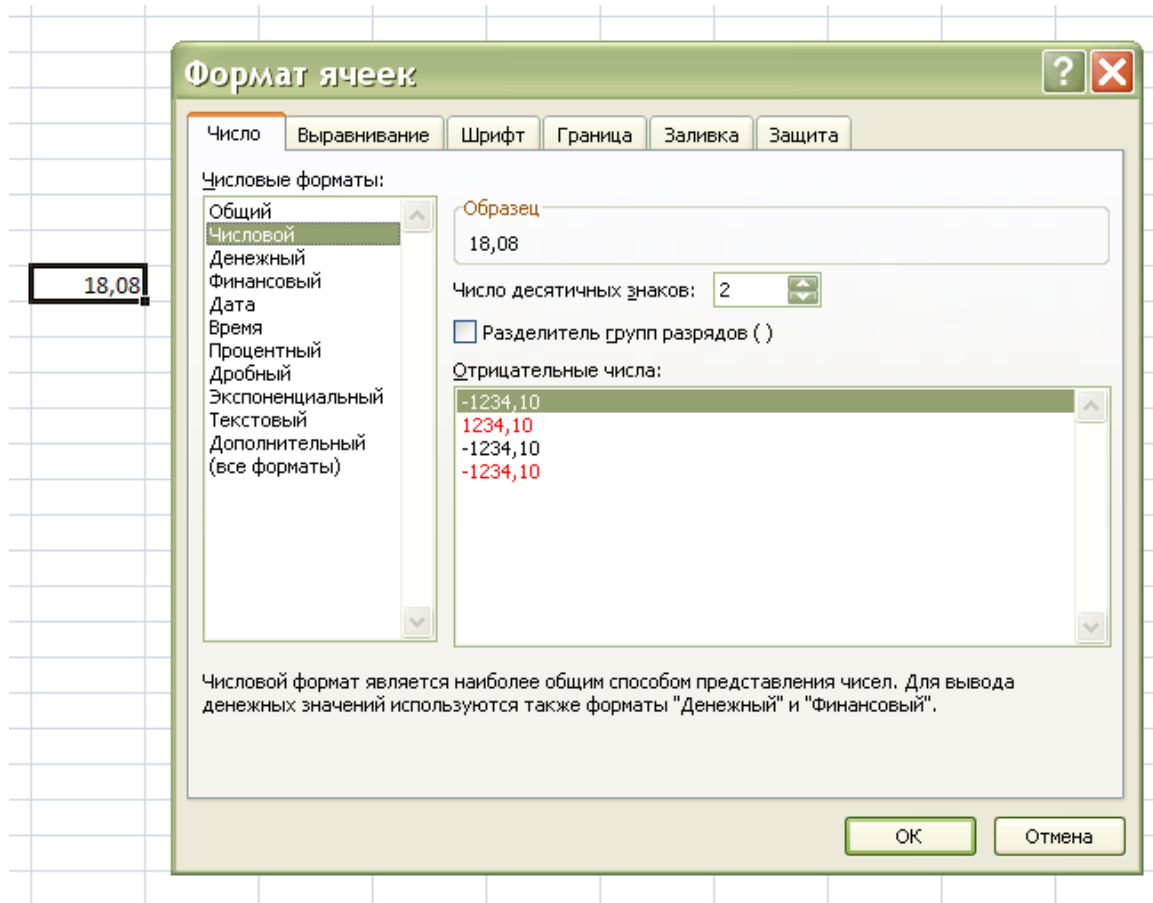
Все комбинации выделяют ту же область,  
но ...

При распространении и **копировании**  
формулы модифицируются по разному

# Формат ячеек

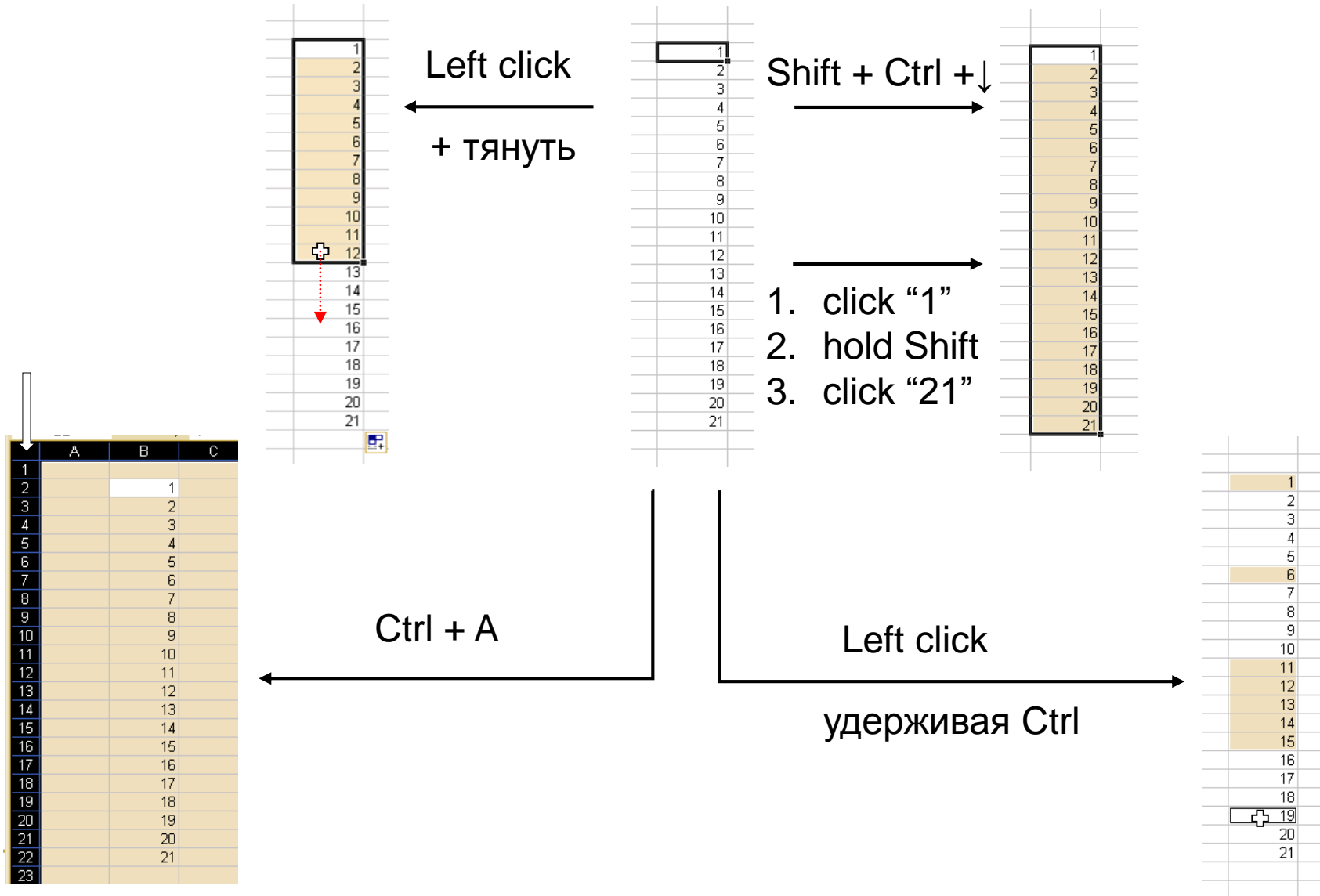
Menu – Format – Cells

Меню- Формат - Ячейка



Right click – Формат ячеек

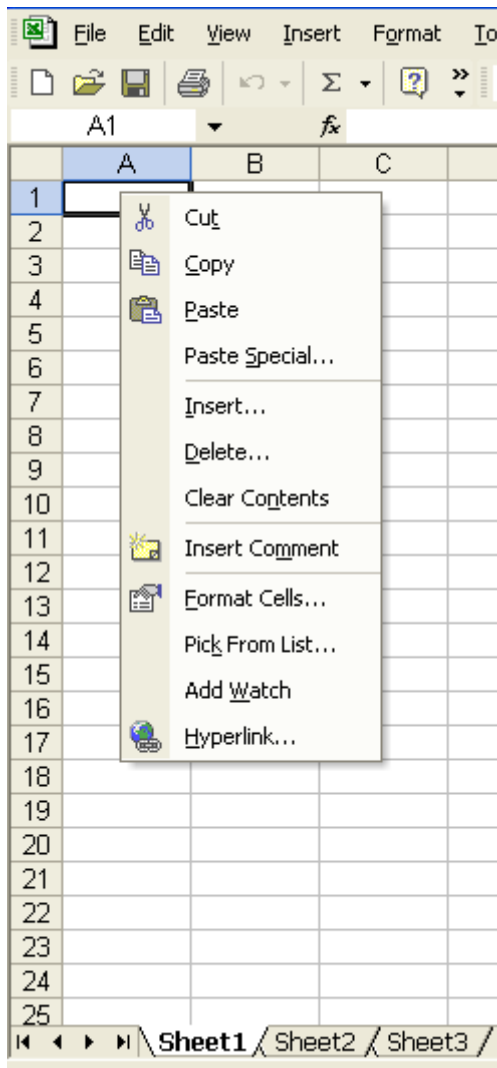
# Выделяем...



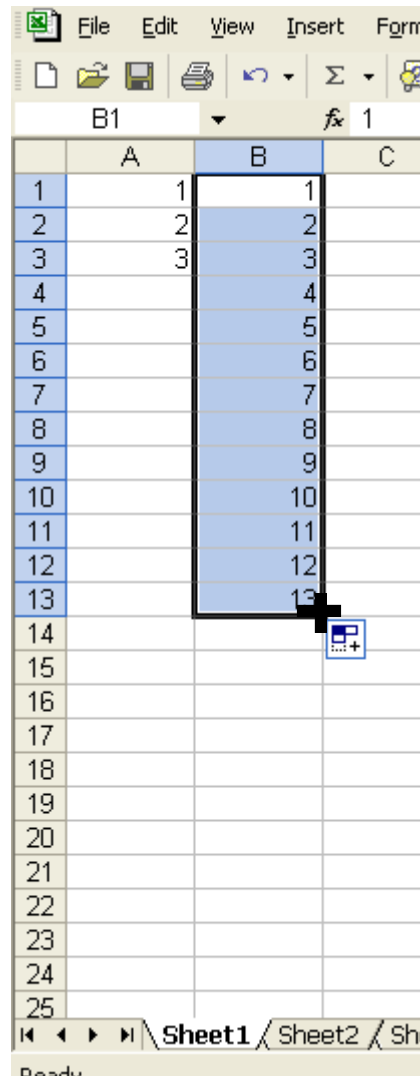


# Ввод данных

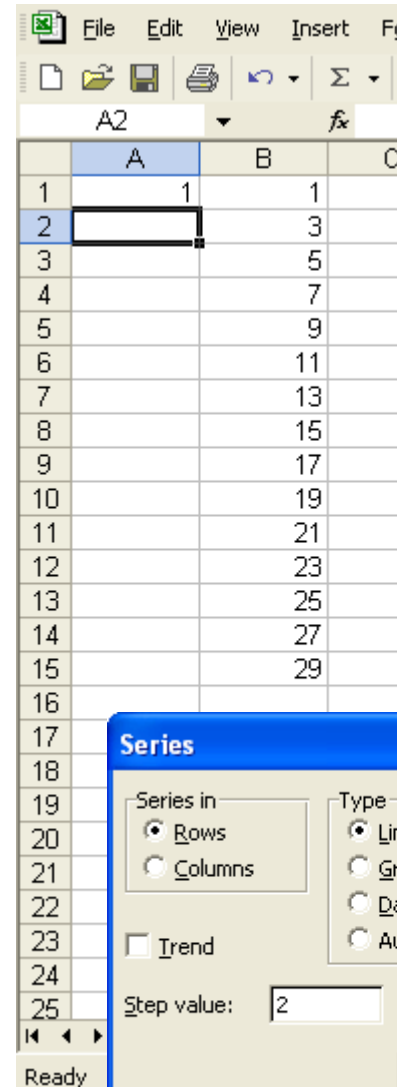
## Меню



## Курсор

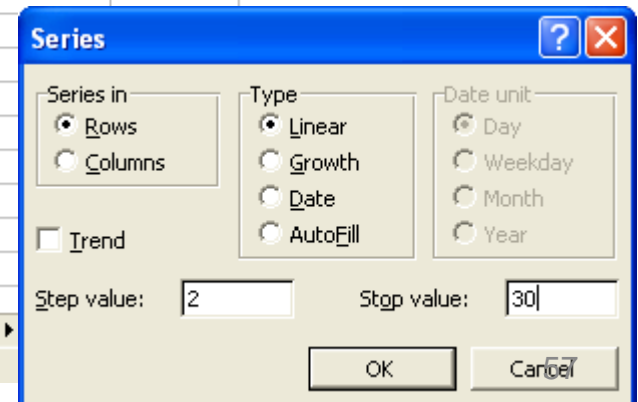


## Прогрессия



Меню –  
Правка –  
Заполнить –  
Прогрессия

Menu – Edit –  
Fill – Series



# закрепление областей

	A	B	C	D	E	F	G	
1		#Integr8_Gene_ID	Primary_Gene_Name	Ordered_Locus_Name	Biological_Names	Start_position	Offset	Direction
2	IGI01006661	secA_1	CT_141	secA	158147	446	F	
3	IGI00976228	CT_610	CT_610		690426	695	R	
4	IGI00976954	nrdB	CT_828	nrdB	977556	1040	F	
5	IGI00972337	fusA	CT_437	fusA	505508	2084	R	
6	IGI00970720	map	CT_851	map	1000369	875	R	
7	IGI00972391	tuf	CT_322	tuf,tu	361980	1184	R	
8	IGI00972153	dnaK	CT_396	dnaK	451614	1982	F	
9	IGI00973083	gap	CT_505	gap,g	583066	1004	R	
10	IGI00977699	rpoD	CT_615	rpoD	694039	1715	R	
11	IGI00971455	clpB	CT_113	clpB	130617	2603	F	
12	IGI00977797	rpsL	CT_439	rpsL,t	508157	371	R	

	A	B	C	D	E	F	G	
1		#Integr8_Gene_ID	Primary_Gene_Name	Ordered_Locus_Name	Biological_Names	Start_position	Offset	Direction
2	IGI01006661	secA_1	CT_141	secA	158147	446	F	
3	IGI00976228	CT_610	CT_610		690426	695	R	
4	IGI00976954	nrdB	CT_828	nrdB	977556	1040	F	
5	IGI00972337	fusA	CT_437	fusA	505508	2084	R	
6	IGI00970720	map	CT_851	map	1000369	875	R	
7	IGI00972391	tuf	CT_322	tuf,tu	361980	1184	R	
8	IGI00972153	dnaK	CT_396	dnaK	451614	1982	F	
9	IGI00973083	gap	CT_505	gap,g	583066	1004	R	
10	IGI00977699	rpoD	CT_615	rpoD	694039	1715	R	
11	IGI00971455	clpB	CT_113	clpB	130617	2603	F	
12	IGI00977797	rpsL	CT_439	rpsL,t	508157	371	R	

	A	R	S	T	U	
1		#Integr8_Gene_ID	Description	EMBL_AC	NCBI_Gis	Inc_prediction
32	IGI00977983	30S ribosomal protein S4	AE00127	3328959	no	
33	IGI00977907	30S ribosomal protein S5	AE00127	3328961	no	
34	IGI00972472	Enolase (EC 4.2.1.11) (2)	AE00127	3329030	no	
35	IGI01006737	Pyruvate Dehydrogenase	AE00127	3328656	no	
36	IGI00974103	Translation initiation factor	AE00127	6578116	no	
37	IGI00978858	Valyl-tRNA synthetase	AE00127	3328717	no	
38	IGI01006633	GTP Binding Protein	AE00127	3328488	no	
39	IGI00977426	50S ribosomal protein L1	AE00127	#####	no	
40	IGI00972350	Elongation factor P1 (E)	AE00127	3328521	no	
41	IGI00970586	Acyl carrier protein (AC)	AE00127	3328645	no	
42	IGI00978368	Superoxide dismutase	AE00127	3328709	no	

выделить границу  
закрепления

Окно → Закрепить  
области

области закреплены  
и видны постоянно  
при прокрутке листа

# Числовая функция

The screenshot displays the Microsoft Excel interface. The active window is titled "Microsoft Excel - Book1". The menu bar includes File, Edit, View, Insert, Format, Tools, Data, Window, and Help. The toolbar shows various icons for file operations and formatting. The formula bar at the top indicates the active cell is B2, containing the formula  $=A2+1$ . The spreadsheet grid shows columns A through M and rows 1 through 31. Cell B2 is selected and contains the numerical value 8. The task pane on the left side of the window lists several tasks, with "Числовая формула" (Number formula) highlighted as the second item. A file dialog box is open in the foreground, showing a "File name:" field and "Files of type:" set to "All Files". A presentation window is partially visible behind the file dialog, showing a slide with the word "текста" (text) and a "Tools" menu.

# Текстовая функция

The screenshot shows the Microsoft Excel interface. The title bar reads "Microsoft Excel - Book1". The menu bar includes "File", "Edit", "View", "Insert", "Format", "Tools", "Data", "Window", and "Help". The toolbar contains various icons for file operations and formatting. The active cell is C3, and the formula bar shows the formula `=CONCATENATE(A3;B3)`. The spreadsheet grid shows the following data:

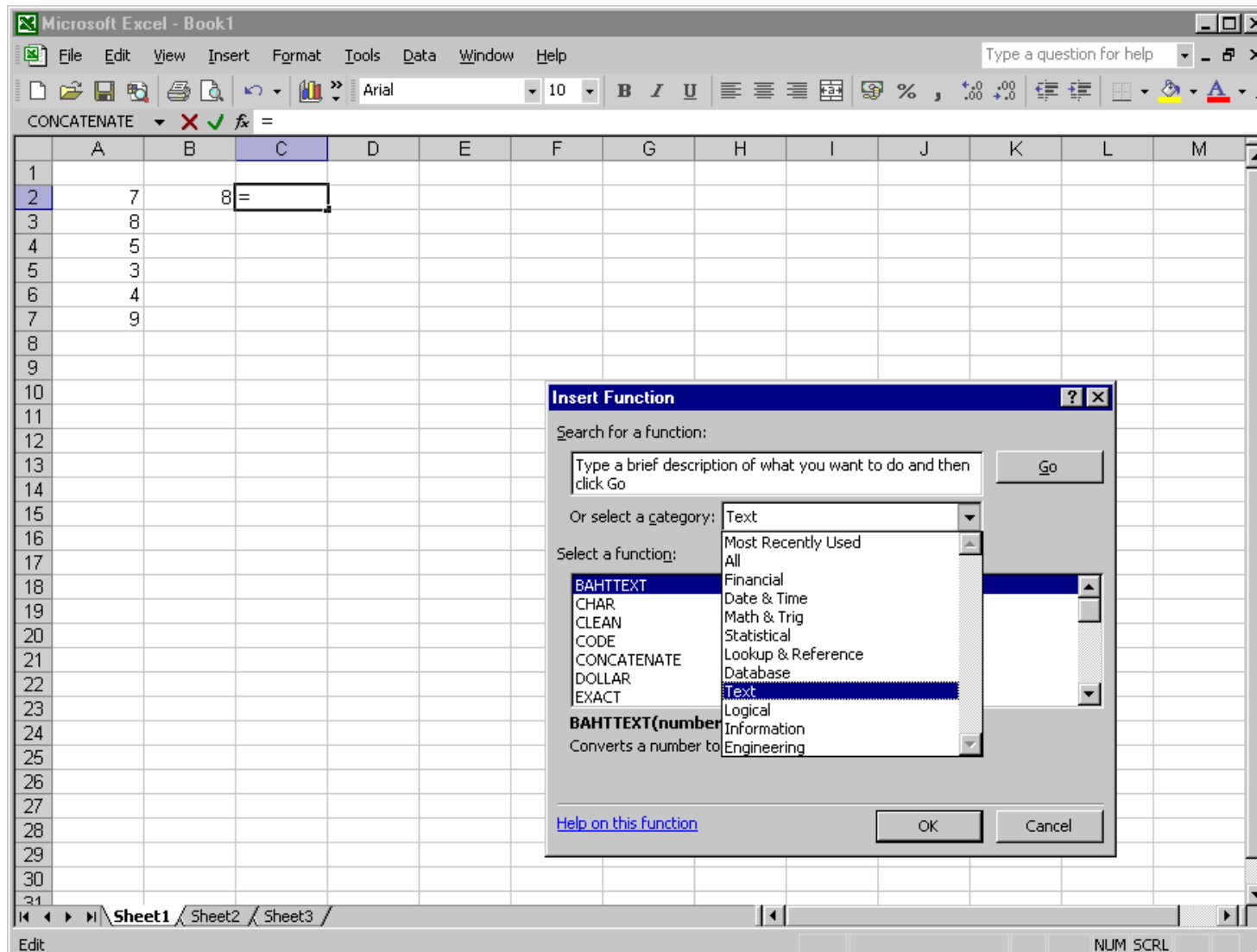
	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2													
3	КОНЬ	ЯК	КОНЬЯК										
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													

The status bar at the bottom shows "Ready" and "NUM SCRL". The sheet tabs at the bottom indicate "Sheet1", "Sheet2", and "Sheet3".

# Мастер функций

Menu – Insert – Function

Меню – Вставка – Функция



# Мастер функций

The screenshot shows the Microsoft Excel interface with the Function Arguments dialog box open for the ROMAN function. The dialog box displays the function name, the number argument (B2), and the resulting Roman numeral (VIII). A yellow lightning bolt icon is positioned over the function name.

Microsoft Excel - Book1

File Edit View Insert Format Tools Data Window Help

Type a question for help

ROMAN =ROMAN(B2)

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2		7	8										
3		8											
4		5											
5		3											
6		4											
7		9											
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													

Function Arguments

ROMAN

Number B2 = 8

Form = number

= "VIII"

Converts an Arabic numeral to Roman, as text.

Number is the Arabic numeral you want to convert.

Formula result = VIII

[Help on this function](#)

OK Cancel

Sheet1 Sheet2 Sheet3

Point NUM SCRL

# Мастер функций

The screenshot shows the Microsoft Excel interface with the 'Function Arguments' dialog box open for the PRODUCT function. The dialog box displays the following information:

- PRODUCT**
- Number1:** A2:A7 = {7;8;5;3;4;9}
- Number2:** = number
- Result:** = 30240
- Description:** Multiplies all the numbers given as arguments.
- Number1:** number1;number2;... are 1 to 30 numbers, logical values, or text representations of numbers that you want to multiply.
- Formula result =** 30240
- [Help on this function](#)
- Buttons:** OK, Cancel

The background spreadsheet shows the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2		7	8 VIII	{A2:A7}									
3		8											
4		5											
5		3											
6		4											
7		9											

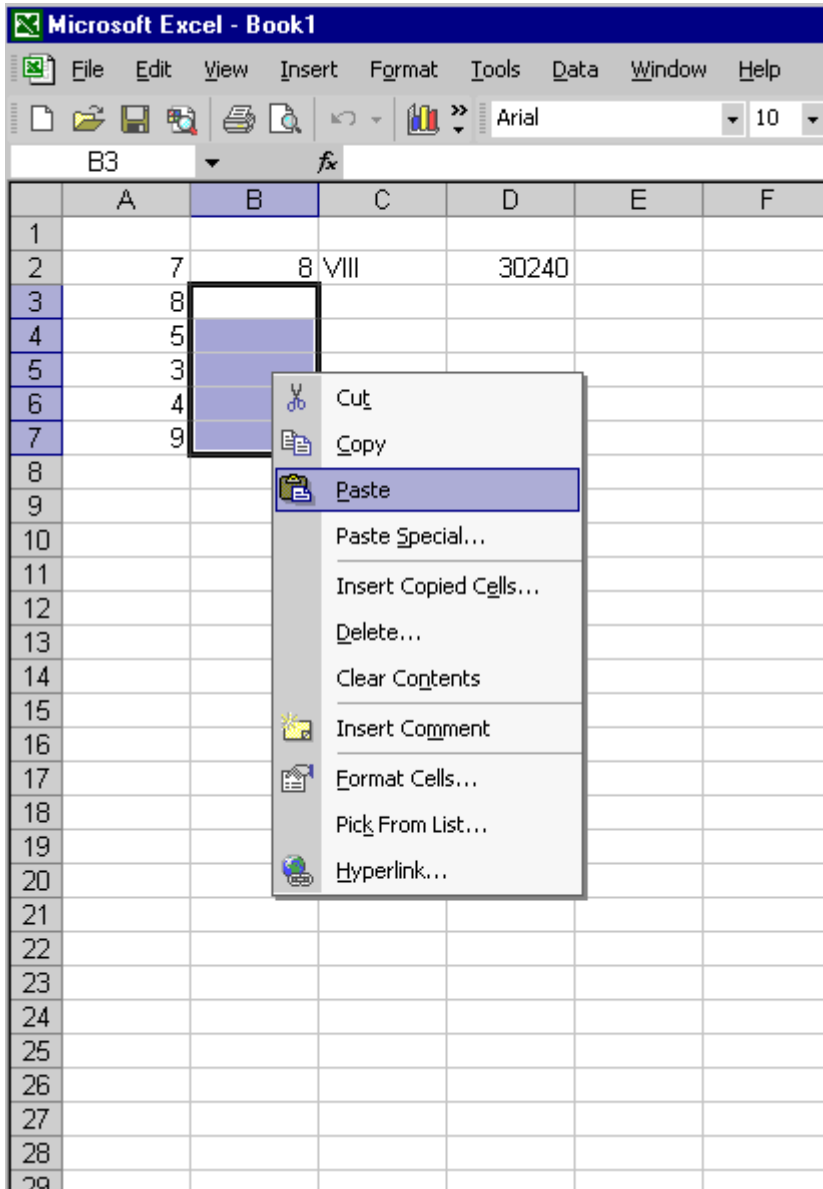
# Копируем формулу...

The screenshot shows the Microsoft Excel interface with the following details:

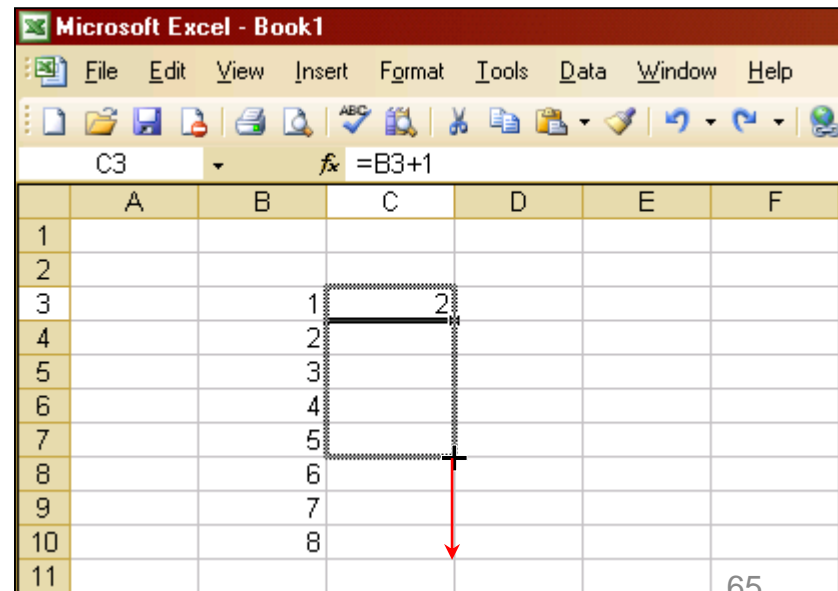
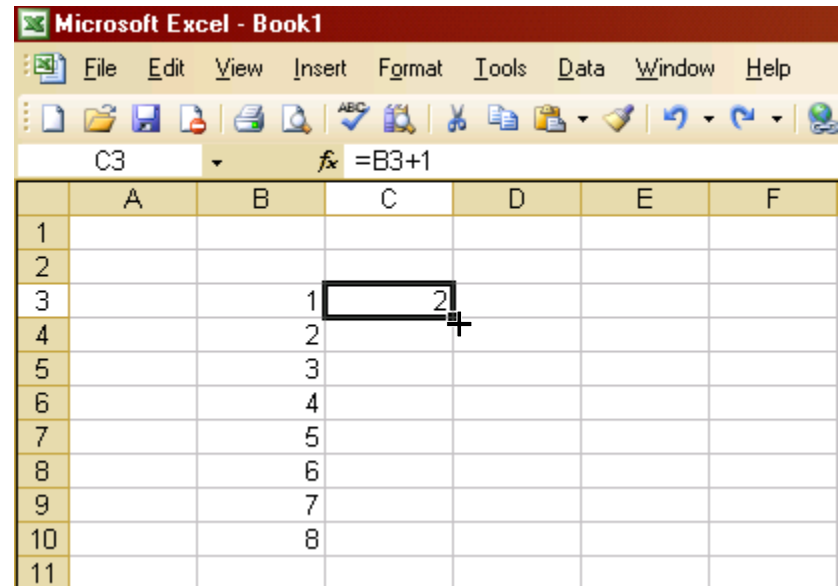
- Window title: Microsoft Excel - Book1
- Menu bar: File, Edit, View, Insert, Format, Tools, Data, Window, Help
- Formula bar: B2, =A2+1
- Grid: Column A contains values 7, 8, 5, 3, 4, 9. Column B is selected. Cell D2 contains the value 30240.
- Context menu (right-click):
  - Cut
  - Copy** (highlighted)
  - Paste
  - Paste Special...
  - Insert...
  - Delete...
  - Clear Contents
  - Insert Comment
  - Format Cells...
  - Pick From List...
  - Add Watch
  - Hyperlink...
- Annotation: An arrow points to the context menu with the text "Меню открывается щелчком правой кнопки мыши".



# ... в другие ячейки



ИЛИ



# Относительные ссылки

The screenshot shows the Microsoft Excel interface with the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2	7	8	VIII	30240									
3	8	9											
4	5	6											
5	3	4											
6	4	5											
7	9	10											
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													

# Относительные ссылки

Microsoft Excel - Book1

File Edit View Insert Format Tools Data Window Help

Type a question for help

Formula bar:  $=B2+1$

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2		7	8	VIII	30240								
3		8	9										
4		5	6										
5		3	4										
6		4	5										
7		9	10										
8													
9													
10													
11													
12													
13													
14													
15													
16													
17													
18													
19													
20													
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													

# Абсолютные и смешанные ссылки

	А	В	С
1			
2			
3			

Первоначальная ссылка

Новая ссылка

**\$A\$1**

абсолютный столбец и абсолютная строка

**\$A\$1**

**A\$1**

относительный столбец и абсолютная строка

**C\$1**

**\$A1**

абсолютный столбец и относительная строка

**\$A3**

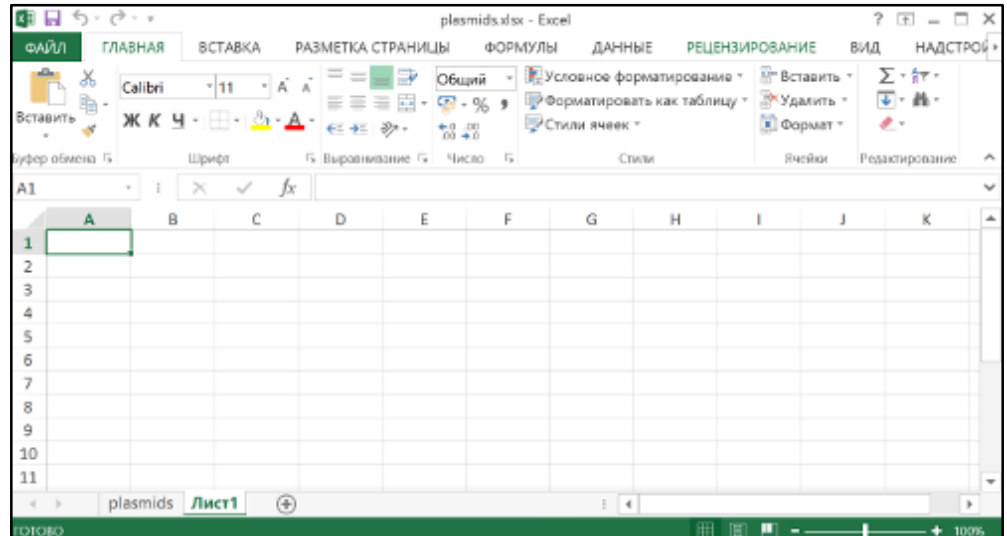
**A1**

относительный столбец и относительная строка

**C3**

# Файл Excel/Google sheet

- Файл Excel имеет расширение .xlsx или .xls
- Файл состоит из листов
- Лист – таблица; состоит из ячеек
- Возможно использование данных из одного .xlsx файла в другом .xlsx файле



# О зачётных навыках

# Задание для выполнения в классе

- Импортировать хромосомную таблицу своей бактерии в Excel
- Могут быть зачтены умения:
  - Фильтр строк и копирование
  - Функция ВПР для связи таблиц
  - Распространение формул
  - Адреса и \$

# Электронные таблицы позволяют

- Обрабатывать данные среднего размера – до нескольких сотен тысяч строк
- Представлять результаты анализа с качеством, приемлемым для публикаций, курсовых и отчетов



# В чем состоит обработка Excel и представление данных?

СПОСОБЫ ОБЛАГОРАЖИВАНИЯ ИЗУМРУДОВ :)

**Excel.** Новейшая, улучшенная разработка фирмы Arthur Groom lab. So

Исходные данные



Отобрать нужные и  
навести порядок



Публикабельный вид



# Питон ПРОТИВ Excel

## Python

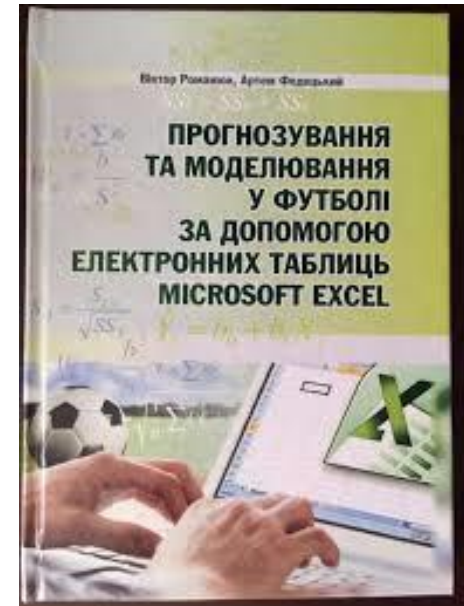


## Guru

## Excel



## dummy



# Питон ПРОТИВ Excel

	python	Excel
Объём данных	ограничен только памятью компьютера	Не более 1млн строк в таблице
Время работы	Можно оптимизировать алгоритм, переписать его на другом языке программирования, C++, например	При 1млн строк даже некоторые простые функции работают часы.
Гибкость	Надо ЗАРАНЕЕ знать или придумать <ul style="list-style-type: none"><li>* алгоритм,</li><li>* параметры, которые можно менять,</li><li>* способ представления результата</li></ul>	Можно быстро менять способ обработки и представления тех же самых данных
Квалификация	Выше секретарши босса но ниже гуру в программировании	От секретарши до квалифицированного биоинформатика
Доступность графических средств	Доступны, но сложнее для освоения и быстрой адаптации	Доступны в ограниченном, но достаточном объёме