SF EDUCATION

Продвинутые навыки работы в Excel (русскоязычный интерфейс)

Текстовые функции

Содержание:

1. Повторения итогов предыдущих частей блока
2. Описание поставленной задачи
3. Вынесение информации по фамилиям
4. Вынесение информации по фамилиям в одной формуле
5. Извлечение информации, находящейся в скобках

Повторения итогов предыдущих частей блока

В продолжение предыдущего занятия, усложняем нашу задачу. На прошлом занятии мы смогли при помощи **простых функций ЛЕВСИМВ и ПРАВСИМВ извлекать из начала и из конца строки определённое количество значений**.

	ДО;АРТИКУЛ;ПРОДУКТ;КОЛ-ВО	ДО	КОЛ-ВО	
2	МС-1; ZVB-762-170-918-FIN; КК; 18	МС-1	18,00	18
3	ЮГ-3; DGS-344-543-979-FIN; ДК; 48	ЮГ-3	48,00	48
4	СБ-3; KJH-112-487-402-FIN; ДП; 47	СБ-3	47,00	47
5	СБ-2; UHD-310-362-461-FIN; НС; 16	СБ-2	16,00	16
6	ЮГ-3; AFS-883-832-420-FIN; ИП; 64	ЮГ-3	64,00	64
7	МС-1; SNH-392-204-635-FIN; КК; 52	МС-1	52,00	52
8	СП-3; KIH-856-361-655-FIN; ДК; 50	СП-3	50,00	50
9	ЮГ-1; NHF-757-809-168-FIN; ДП; 14	ЮГ-1	14,00	14
10	ЮГ-3; DBG-338-116-638-FIN; КК; 18	ЮГ-3	18,00	18
11	СБ-1; UGD-674-193-770-FIN; ДК; 60	СБ-1	60,00	60
12	МС-1; NST-204-148-417-FIN; ДП; 46	МС-1	46,00	46
13	СБ-2; AYH-674-849-158-FIN; КК; 22	СБ-2	22,00	22
14	МС-2; DHY-568-730-791-FIN; ДК; 46	МС-2	46,00	46
15	ЮГ-2; JSG-783-175-549-FIN; ДП; 28	ЮГ-2	28,00	28
16	СБ-3; LDY-155-659-980-FIN; НС; 59	СБ-3	59,00	59
17			588,00	588

Также, с помощью **комбинаций функций ЛЕВСИМВ- ПОИСК и ПРАВСИМВ-ДЛСТР-ПОИСК** смогли **извлечь из текстовой строки** вне зависимости от того или иного количества букв либо цифр, до или после определённого знака, который их разделяет, **цифровое значение**.

19	Усложнение		
20	АРТИКУЛ	Букв.	Кол-ое
21	ZVBS_76	ZVBS	76
22	HSGSFS_34454	HSGSFS	34454
23	KJ_1	KJ	1
24	UHAD_3105	UHAD	3105
25	AFS_883	AFS	883
26	SNH_392	SNH	392
27	K_856555	K	856555
28	NF_7575	NF	7575
29	DBG_3	DBG	3
30	UGD_6	UGD	6
31	NSDDT_2044	NSDDT	2044
32	AH_67	AH	67
33	DY_56	DY	56
34	J_78	J	78
35	LDY_15	LDY	15

Описание поставленной задачи

Усложняем нашу задачу. Представим, что у нас есть следующая выгрузка:

37	Усложнение
38	Данные по сотрудникам
39	МС202205_Кузьмина-ANB1982001(95428)
40	ЮГ202201_Беспалов-LDN2104167(65633)
41	СБ202202_Богданова-BNL8163819(29456)
42	СБ202201_Калинин-BNL4000377(40848)
43	ЮГ202204_Иванова-LDN5824771(62783)
44	МС202201_Лаврентьева-ANB3611381(85180)
45	СП202210_Софронов-YDH4488449(89414)
46	ЮГ202201_Назарова-LDN4341377(29076)
47	ЮГ202205_Савельева-LDN2118346(82063)
48	СБ202204_Карасев-BNL8927104(91958)
49	МС202204_Иванова-ANB5254736(10252)
50	СБ202202_Андреева-BNL8788852(56471)
51	МС202212_Трофимов-ANB1271678(25683)
52	ЮГ202207_Лопатина-LDN7327397(84370)
53	СБ202201_Глушков-BNL6476274(60149)

Изучив эту выгрузку, становится понятно, что скорее всего **это обозначение неких продаж из какой-то программы**. Здесь чётко видны:

- Регион продажи;
- Год;
- Месяц;
- Фамилия, наверное, менеджера;
- Некое обозначение, скорее всего либо **табельный номер сотрудника**, либо **идентификационный номер** сотрудника, либо идентификационный номер точки продажи;

- **В скобках** – обозначение некоего **набора цифр**, скорее всего, это или премия, или зарплата, или какой-то общий доход.

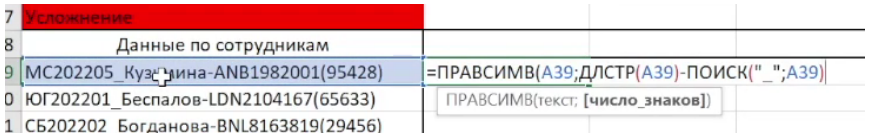
В рамках нашей текущей задачи нам **необходимо извлечь информацию по фамилии сотрудника-менеджера и информацию по цифровому набору**, а именно: по количеству цифр, которое стоит **между скобками**.

Разбираемся, как можно извлечь информацию с помощью формул и поможет нам знание, как делать **комбинации между двумя текстовыми функциями**.

Вынесение информации по фамилиям

Сделаем это поэтапно.

Мы знаем, что с помощью комбинации **функции ПРАВСИМВ, ДЛСТР и ПОИСК** сможем извлечь всю ту информацию, что идёт после фамилии.



Делаем это:

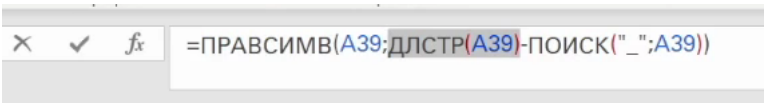


37	Усложнение	
38	Данные по сотрудникам	
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина-ANB1982001(95428)
40	ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов-LDN2104167(65633)
41	СБ202202_Богданова-BNL8163819(29456)	Богданова-BNL8163819(29456)
42	СБ202201_Калинин-BNL4000377(40848)	Калинин-BNL4000377(40848)
43	ЮГ202204_Иванова-LDN5824771(62783)	Иванова-LDN5824771(62783)
44	МС202201_Лаврентьева-ANB3611381(85180)	Лаврентьева-ANB3611381(85180)
45	СП202210_Софронов-YDH4488449(89414)	Софронов-YDH4488449(89414)
46	ЮГ202201_Назарова-LDN4341377(29076)	Назарова-LDN4341377(29076)
47	ЮГ202205_Савельева-LDN2118346(82063)	Савельева-LDN2118346(82063)
48	СБ202204_Карасев-BNL8927104(91958)	Карасев-BNL8927104(91958)
49	МС202204_Иванова-ANB5254736(10252)	Иванова-ANB5254736(10252)
50	СБ202202_Андреева-BNL8788852(56471)	Андреева-BNL8788852(56471)
51	МС202212_Трофимов-ANB1271678(25683)	Трофимов-ANB1271678(25683)
52	ЮГ202207_Лопатина-LDN7327397(84370)	Лопатина-LDN7327397(84370)
53	СБ202201_Глушков-BNL6476274(60149)	Глушков-BNL6476274(60149)
54		

Проверяем. Да, с помощью этих функций всё получилось.

Мы **можем прочитать эту формулу**:

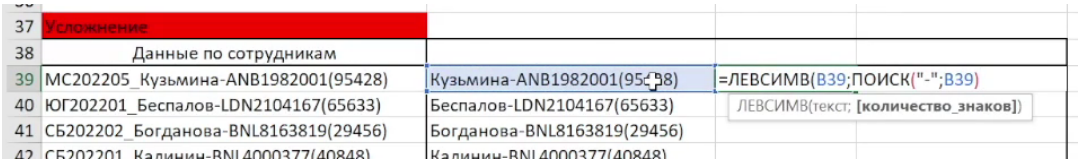
- Из наших ячеек с текстом, с набором буквенных значений и количественных цифр, с учётом формулы, полученной из **функции ДЛСТР** вычитаем информацию с пометкой «извлеки нам из ячейки A39 всю информацию, которая продолжается после данного разделительного знака».



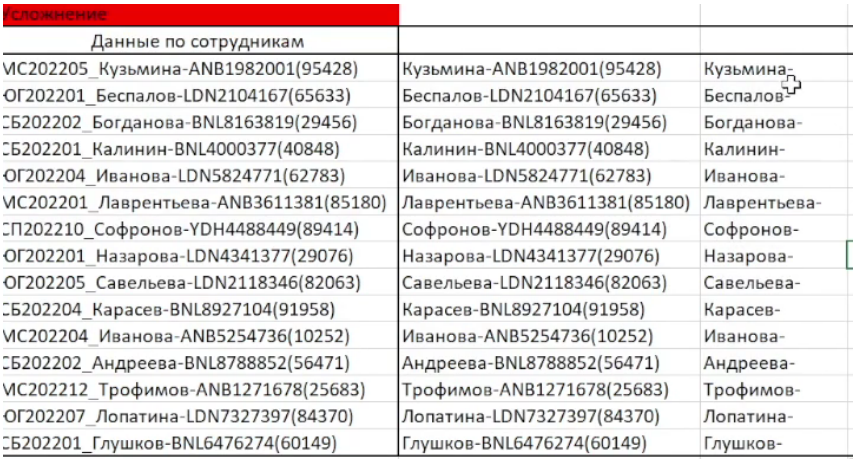
В итоге мы получили, что хотели.

Теперь из этого нужно **получить информацию по фамилиям сотрудников**. Применяем знания прошлого занятия.

Поскольку мы уже сделали итерацию извлечения необходимой нам информации, то можно **ссылаться на полученную ячейку** для извлечения информации по фамилии сотрудника.



Здесь **разделительным знаком** ставим **«минус»** либо «**тире**».



Усложнение		
Данные по сотрудникам		
МС202205_Кузьмина-ANB1982001(95428)	Кузьмина-ANB1982001(95428)	Кузьмина-
ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов-LDN2104167(65633)	Беспалов
СБ202202_Богданова-BNL8163819(29456)	Богданова-BNL8163819(29456)	Богданова-
СБ202201_Калинин-BNL4000377(40848)	Калинин-BNL4000377(40848)	Калинин-
ЮГ202204_Иванова-LDN5824771(62783)	Иванова-LDN5824771(62783)	Иванова-
МС202201_Лаврентьева-ANB3611381(85180)	Лаврентьева-ANB3611381(85180)	Лаврентьева-
СП202210_Софронов-YDH4488449(89414)	Софронов-YDH4488449(89414)	Софронов-
ЮГ202201_Назарова-LDN4341377(29076)	Назарова-LDN4341377(29076)	Назарова-
ЮГ202205_Савельева-LDN2118346(82063)	Савельева-LDN2118346(82063)	Савельева-
СБ202204_Карасев-BNL8927104(91958)	Карасев-BNL8927104(91958)	Карасев-
МС202204_Иванова-ANB5254736(10252)	Иванова-ANB5254736(10252)	Иванова-
СБ202202_Андреева-BNL8788852(56471)	Андреева-BNL8788852(56471)	Андреева-
МС202212_Трофимов-ANB1271678(25683)	Трофимов-ANB1271678(25683)	Трофимов-
ЮГ202207_Лопатина-LDN7327397(84370)	Лопатина-LDN7327397(84370)	Лопатина-
СБ202201_Глушков-BNL6476274(60149)	Глушков-BNL6476274(60149)	Глушков-

Мы получили необходимый результат, который должны были получить, но забыли отнять один знак:

```
=ЛЕВСИМВ(B39;ПОИСК("-";B39)-1)
```

Итог:

38	Данные по сотрудникам		
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина-ANB1982001(95428)	Кузьмина
40	ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов-LDN2104167(65633)	Беспалов
41	СБ202202_Богданова-BNL8163819(29456)	Богданова-BNL8163819(29456)	Богданова
42	СБ202201_Калинин-BNL4000377(40848)	Калинин-BNL4000377(40848)	Калинин
43	ЮГ202204_Иванова-LDN5824771(62783)	Иванова-LDN5824771(62783)	Иванова
44	МС202201_Лаврентьева-ANB3611381(85180)	Лаврентьева-ANB3611381(85180)	Лаврентьева
45	СП202210_Софронов-YDH4488449(89414)	Софронов-YDH4488449(89414)	Софронов
46	ЮГ202201_Назарова-LDN4341377(29076)	Назарова-LDN4341377(29076)	Назарова
47	ЮГ202205_Савельева-LDN2118346(82063)	Савельева-LDN2118346(82063)	Савельева
48	СБ202204_Карасев-BNL8927104(91958)	Карасев-BNL8927104(91958)	Карасев
49	МС202204_Иванова-ANB5254736(10252)	Иванова-ANB5254736(10252)	Иванова
50	СБ202202_Андреева-BNL8788852(56471)	Андреева-BNL8788852(56471)	Андреева
51	МС202212_Трофимов-ANB1271678(25683)	Трофимов-ANB1271678(25683)	Трофимов
52	ЮГ202207_Лопатина-LDN7327397(84370)	Лопатина-LDN7327397(84370)	Лопатина
53	СБ202201_Глушков-BNL6476274(60149)	Глушков-BNL6476274(60149)	Глушков

Таким образом, зная и понимая **алгоритм извлечения с помощью комбинаций функций**, с помощью двух итераций мы смогли извлечь необходимую информацию.

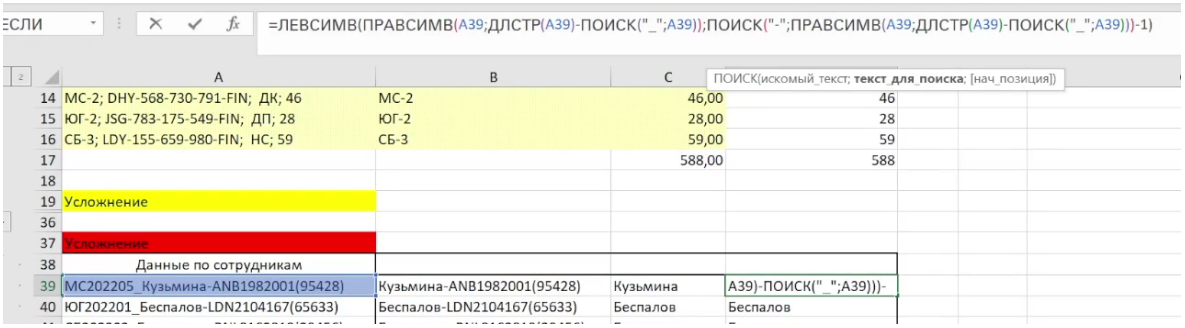
Вынесение информации по фамилиям в одной формуле

Изначально мы делали **одну итерацию**, где извлекали информацию после «**нижнего подчёркивания**» и уже из полученного результата извлекали фамилию сотрудника **до разделительного знака** «**тире**». И всё время ссылались на промежуточный результат.

По факту мы можем скопировать данную формулу

```
=ПРАВСИМВ(A39;ДЛСТР(A39)-ПОИСК("_";A39))
```

И подставить, то есть **заменить этой формулой искомый текст.**



У нас получился идентичный результат:

38	Данные по сотрудникам			
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина-ANB1982001(95428)	Кузьмина	Кузьмина
40	ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов-LDN2104167(65633)	Беспалов	Беспалов
41	СБ202202_Богданова-BNL8163819(29456)	Богданова-BNL8163819(29456)	Богданова	Богданова
42	СБ202201_Калинин-BNL4000377(40848)	Калинин-BNL4000377(40848)	Калинин	Калинин
43	ЮГ202204_Иванова-LDN5824771(62783)	Иванова-LDN5824771(62783)	Иванова	Иванова
44	МС202201_Лаврентьева-ANB3611381(85180)	Лаврентьева-ANB3611381(85180)	Лаврентьева	Лаврентьева
45	СП202210_Софронов-YDH4488449(89414)	Софронов-YDH4488449(89414)	Софронов	Софронов
46	ЮГ202201_Назарова-LDN4341377(29076)	Назарова-LDN4341377(29076)	Назарова	Назарова
47	ЮГ202205_Савельева-LDN2118346(82063)	Савельева-LDN2118346(82063)	Савельева	Савельева
48	СБ202204_Карасев-BNL8927104(91958)	Карасев-BNL8927104(91958)	Карасев	Карасев
49	МС202204_Иванова-ANB5254736(10252)	Иванова-ANB5254736(10252)	Иванова	Иванова
50	СБ202202_Андреева-BNL8788852(56471)	Андреева-BNL8788852(56471)	Андреева	Андреева
51	МС202212_Трофимов-ANB1271678(25683)	Трофимов-ANB1271678(25683)	Трофимов	Трофимов
52	ЮГ202207_Лопатина-LDN7327397(84370)	Лопатина-LDN7327397(84370)	Лопатина	Лопатина
53	СБ202201_Глушков-BNL6476274(60149)	Глушков-BNL6476274(60149)	Глушков	Глушков
54				

Только уже **в одной формуле**.

Рекомендация. Если понятно, что формула будет громоздкой и сложной для понимания логики её написания, то нужно стараться делить такую формулу на определённые пункты расчёта – итерации расчёта.

Извлечение информации, находящейся в скобках

По сути, это не составляет проблемы. Убираем лишнее и именуем столбики.

	Данные по сотрудникам	ФИО	Премия
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина	
40	ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов	
41	СБ202202_Богданова-BNL8163819(29456)	Богданова	
42	СБ202201_Калинин-BNL4000377(40848)	Калинин	
43	ЮГ202204_Иванова-LDN5824771(62783)	Иванова	
44	МС202201_Лаврентьева-ANB3611381(85180)	Лаврентьева	
45	СП202210_Софронов-YDH4488449(89414)	Софронов	
46	ЮГ202201_Назарова-LDN4341377(29076)	Назарова	
47	ЮГ202205_Савельева-LDN2118346(82063)	Савельева	
48	СБ202204_Карасев-BNL8927104(91958)	Карасев	
49	МС202204_Иванова-ANB5254736(10252)	Иванова	
50	СБ202202_Андреева-BNL8788852(56471)	Андреева	
51	МС202212_Трофимов-ANB1271678(25683)	Трофимов	
52	ЮГ202207_Лопатина-LDN7327397(84370)	Лопатина	
53	СБ202201_Глушков-BNL6476274(60149)	Глушков	

Мы знаем, **как извлекать информацию с помощью определённых знаков разделения**. Любой знак разделения для нас является триггером для вписывания его в формулы текстовых функций.

Если мы видим, что данные однотипны, то есть имеется **идентичное количество в наборах знаков**, то извлечь информацию можно с помощью **ПРАВСИМВ.**

Но у нас помеха – закрывающаяся скобка в конце.

Как один из вариантов, можно:

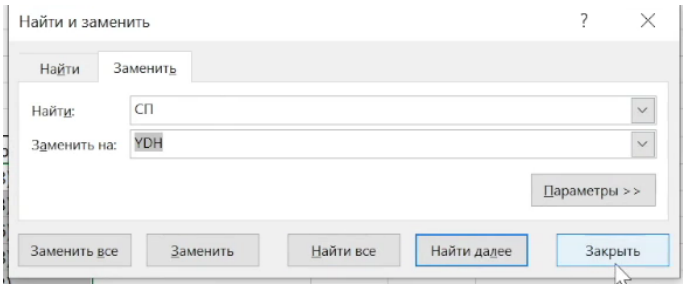
- С помощью **ПСТР**, сослаться на текст;
- Найти изначально открывающуюся скобку;
- Далее есть закрывающаяся скобка.

Данные по сотрудникам	ФИО	Премия
МС202205_Кузьмина-ANB1982001(95428)	Кузьмина	=ПСТР(A39;ПОИСК("(";A39);ПОИСК(")";A39))
ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов	

И есть полученный результат:

38	Данные по сотрудникам	ФИО	Премия
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина	(95428)
40	ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов	(65633)
41	СБ202202_Богданова-BNL8163819(29456)	Богданова	(29456)
42	СБ202201_Калинин-BNL4000377(40848)	Калинин	(40848)
43	ЮГ202204_Иванова-LDN5824771(62783)	Иванова	(62783)
44	МС202201_Лаврентьева-ANB3611381(85180)	Лаврентьева	(85180)
45	СП202210_Софронов-YDH4488449(89414)	Софронов	(89414)
46	ЮГ202201_Назарова-LDN4341377(29076)	Назарова	(29076)
47	ЮГ202205_Савельева-LDN2118346(82063)	Савельева	(82063)
48	СБ202204_Карасев-BNL8927104(91958)	Карасев	(91958)
49	МС202204_Иванова-ANB5254736(10252)	Иванова	(10252)
50	СБ202202_Андреева-BNL8788852(56471)	Андреева	(56471)
51	МС202212_Трофимов-ANB1271678(25683)	Трофимов	(25683)
52	ЮГ202207_Лопатина-LDN7327397(84370)	Лопатина	(84370)
53	СБ202201_Глушков-BNL6476274(60149)	Глушков	(60149)

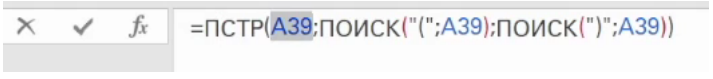
Но это **текстовый формат**, и мы с ним не сможем работать. Можно было сделать через **ctrlF** заменить это всё – скопировать, вставить как значение и потом произвести замены, отформатировать тип данных.



Но это слишком много ненужных итераций.

Чтобы лучше понять, **прочтём ещё раз нашу формулу**:

- Мы из извлекаемого текста высчитываем начальную позицию с обозначением «открывающие скобки» и говорим: «извлеки нам то количество знаков, которое будет между открывшейся скобкой и закрывшейся скобкой».



Исходя из этого, можно просто сказать: **«+ один знак»**. Мы от открывающейся скобки говорим: «**прибавь еще один знак**», то есть «**исключи этот знак**» и имеется ввиду «**открывающаяся скобка**».

Данные по сотрудникам	ФИО	Премия
МС202205_Кузьмина-ANB1982001(95428)	Кузьмина	+1;ПОИСК(")";
ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов	(65633)

Он исключил.

	Данные по сотрудникам	ФИО	Премия
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина	95428)
40	ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов	65633)
41	СБ202202_Богданова-BNL8163819(29456)	Богданова	29456)
42	СБ202201_Калинин-BNL4000377(40848)	Калинин	40848)
43	ЮГ202204_Иванова-LDN5824771(62783)	Иванова	62783)
44	МС202201_Лаврентьева-ANB3611381(85180)	Лаврентьева	85180)
45	СП202210_Софронов-YDH4488449(89414)	Софронов	89414)
46	ЮГ202201_Назарова-LDN4341377(29076)	Назарова	29076)
47	ЮГ202205_Савельева-LDN2118346(82063)	Савельева	82063)
48	СБ202204_Карасев-BNL8927104(91958)	Карасев	91958)
49	МС202204_Иванова-ANB5254736(10252)	Иванова	10252)
50	СБ202202_Андреева-BNL8788852(56471)	Андреева	56471)
51	МС202212_Трофимов-ANB1271678(25683)	Трофимов	25683)
52	ЮГ202207_Лопатина-LDN7327397(84370)	Лопатина	84370)
53	СБ202201_Глушков-BNL6476274(60149)	Глушков	60149)

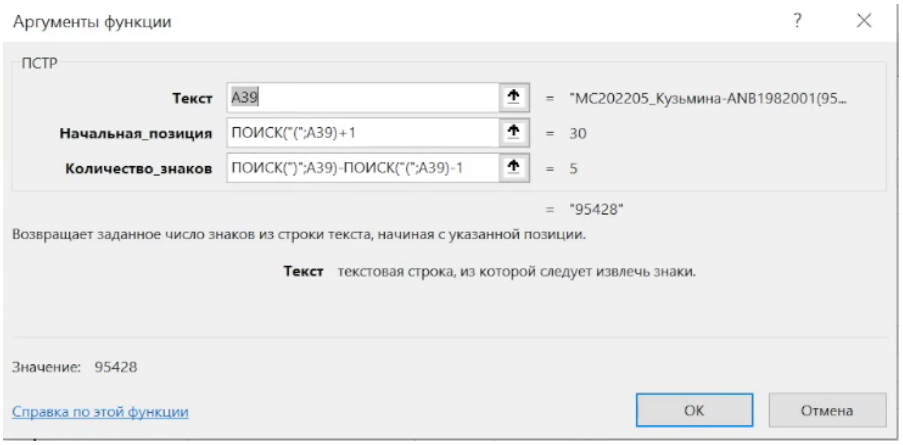
Теперь нужно, чтобы **закрывающаяся скобка тоже исчезла**. Делаем это следующим образом:

38	Данные по сотрудникам	ФИО	Премия
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина	A39)-ПОИСК("(")
40	ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов	65633)

И мы получили необходимый результат в рамках нашей задачи:

38	Данные по сотрудникам	ФИО	Премия
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина	95428
40	ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов	65633
41	СБ202202_Богданова-BNL8163819(29456)	Богданова	29456
42	СБ202201_Калинин-BNL4000377(40848)	Калинин	40848
43	ЮГ202204_Иванова-LDN5824771(62783)	Иванова	62783
44	МС202201_Лаврентьева-ANB3611381(85180)	Лаврентьева	85180
45	СП202210_Софронов-YDH4488449(89414)	Софронов	89414
46	ЮГ202201_Назарова-LDN4341377(29076)	Назарова	29076
47	ЮГ202205_Савельева-LDN2118346(82063)	Савельева	82063
48	СБ202204_Карасев-BNL8927104(91958)	Карасев	91958
49	МС202204_Иванова-ANB5254736(10252)	Иванова	10252
50	СБ202202_Андреева-BNL8788852(56471)	Андреева	56471
51	МС202212_Трофимов-ANB1271678(25683)	Трофимов	25683
52	ЮГ202207_Лопатина-LDN7327397(84370)	Лопатина	84370
53	СБ202201_Глушков-BNL6476274(60149)	Глушков	60149

Не забываем, что если нажать на функцию, то **данные цифры будут отражены, как текстовый тип данных**.



Прибавляем 0:

38	Данные по сотрудникам	ФИО	Премия
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина	A39)-1)+0
40	ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов	65633

Проверяем, всё ли работает.

38	Данные по сотрудникам	ФИО	Премия
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина	95428
40	ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов	65633
41	СБ202202_Богданова-BNL8163819(29456)	Богданова	29456
42	СБ202201_Калинин-BNL4000377(40848)	Калинин	40848
43	ЮГ202204_Иванова-LDN5824771(62783)	Иванова	62783
44	МС202201_Лаврентьева-ANB3611381(85180)	Лаврентьева	85180
45	СП202210_Софронов-YDH4488449(89414)	Софронов	89414
46	ЮГ202201_Назарова-LDN4341377(29076)	Назарова	29076
47	ЮГ202205_Савельева-LDN2118346(82063)	Савельева	82063
48	СБ202204_Карасев-BNL8927104(91958)	Карасев	91958
49	МС202204_Иванова-ANB5254736(10252)	Иванова	10252
50	СБ202202_Андреева-BNL8788852(56471)	Андреева	56471
51	МС202212_Трофимов-ANB1271678(25683)	Трофимов	25683
52	ЮГ202207_Лопатина-LDN7327397(84370)	Лопатина	84370
53	СБ202201_Глушков-BNL6476274(60149)	Глушков	60149
54			908764

Всё работает корректно.

Ещё раз прочтём формулу:



- У нас есть искомый текст и мы всё время ссылаемся на изначальный источник данных;
- Далее мы говорим: «извлеки нам информацию из данной ячейки, триггером для которой является «**начальная позиция +1**», то есть «**не включая первую искомую позицию**»;
- Количество знаков для тебя будет обозначать то, что **закрывающая скобка является окончанием** того количества набора знаков, которое в этих скобках присутствовало;
- При этом, отними информацию.

Мы прибавляем для того, чтобы это не вносить в отражение информации в ячейки, а в конце мы минусуем, чтобы нам с конца это удалить. Но при этом, понимаем, что **поиск открывающейся скобки – это не тоже самое, что закрывающейся скобки.**

То есть, если бы мы сделали закрывающуюся скобку:



38	Данные по сотрудникам	ФИО	Премия
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина	A39)-ПОИСК(")";

Он бы нас не понял:



38	Данные по сотрудникам	ФИО	Премия
39	МС202205_Кузьмина-ANB1982001(95428)	Кузьмина	#ЗНАЧ!
40	ЮГ202201_Беспалов-LDN2104167(65633)	Беспалов	65633

Потому, что у нас значений после скобки нет.

Теперь мы можем **с помощью формул извлекать любое значение из текстового формата типа данных** при выгрузке либо при наборе – неважно с каким исходником данных приходится работать.

С помощью **комбинаций текстовых функций мы можем извлекать любое значение,** которое нам необходимо.