

Limpieza de Base

- Sólo se tomará en cuenta la información de las carreras del Departamento de Matemáticas, es decir, Actuaría, Matemáticas, Matemáticas Aplicadas y Ciencias de la computación.
- Para encontrar los horarios de cada materia se observó que la estructura que se sigue es la siguiente:

“<http://www.fcencias.unam.mx/docencia/horarios/a/b/c>”

De esta manera se hicieron 3 ciclos de tipo “for”, inicialmente los valores de a , b y c eran:

a : 20081, 20082, 20091, 20092, 20101, 20102, \dots , 20201

b : 1 : 10000

c : 1 : 10000

- Se buscó un patrón en las páginas para disminuir el número de combinaciones por lo que los valores que se encontraron fueron: a = semestre, b = plan de estudios, c = número de materia, por lo que los ciclos corren:

Semestre: 20081, 20082, 20091, 20092, 20101, 20102, \dots , 20201

Plan: Planes de estudio

Núm. Materia: 1 : 10000

Los planes de estudio de cada carrera son:

* Actuaría:

1972 \rightarrow plan = 214

2000 \rightarrow plan = 119

2006 \rightarrow plan = 1176

2015 \rightarrow plan = 2017

Los planes vigentes son los últimos tres.

* Ciencias de la Computación:

1994 \rightarrow plan = 218

2013 \rightarrow plan = 1556

* Matemáticas:

1983 \rightarrow plan = 217

* Matemáticas Aplicadas:

2017 \rightarrow plan = 2055

PLAN	CLAVE
Actuaría	
1972	214
2000	119
2006	1176
2015	2017
Ciencias de la Computación	
1994	218
2013	1556
Matemáticas	
1983	217
Matemáticas Aplicadas	
2017	2055

- Al inicio se encontraron tres tipos de grupos de las páginas de horarios de la FC:

- a) Las páginas correspondientes al semestre 2020 – 1, la cual no contiene la información de los exámenes finales. <http://www.fciencias.unam.mx/docencia/horarios/20201/2017/1707>

Grupo 9301, 129 lugares, 84 alumnos
Profesor José Luis Navarro Urrutia lu mi vi 13 a 14 Aula Magna I
Ayudante Luz Candy Becerril Palacios ma ju 13 a 14 Aula Magna I
Ayudante Gabriela Yaneth Romo Cordoba ma ju 13 a 14
Ayudante Adrián Gallardo Pacheco ma ju 13 a 14

- b) Páginas correspondientes a semestres entre el 2018–2 y el semestre 2019 – 2, las cuales tienen información del número de lugares disponibles por salón. <http://www.fciencias.unam.mx/docencia/horarios/20182/217/625>

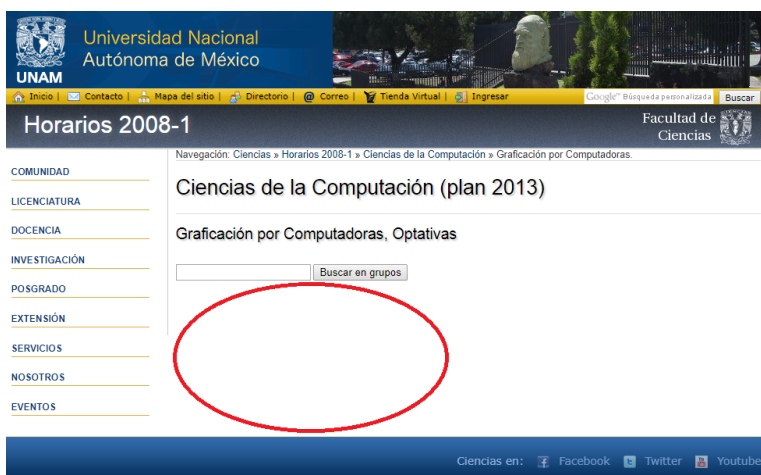
Grupo 9027, 112 lugares, 68 alumnos
Exámenes finales martes 29 de mayo 2018 y martes 5 de junio 2018 de 18 a 20
 Profesor Martín Martínez Estrada lu mi vi 18 a 19 Aula Magna I
 Ayudante Eleazar Bello Cervantes ma ju 18 a 19 Aula Magna I
 Ayudante José Eduardo Quintero García ma ju 18 a 19
 Presentación

- c) Páginas correspondientes a semestres anteriores al 2018 – 1, incluyéndolo, éstas páginas no contienen la información del número de lugares disponibles por cada salón. <http://www.fciencias.unam.mx/docencia/horarios/20181/1556/399>

Grupo 9259 72 alumnos
Exámenes finales jueves 11 de enero 2018 y jueves 18 de enero 2018 de 18 a 20.
 Profesor Francisco Sánchez Villarreal lu mi vi 18 a 19 P213
 Ayudante Santiago Lara Jiménez ma ju 18 a 19 P213
 Ayudante José Oscar Rosales Vergara ma ju 18 a 19

- Al querer obtener la información de las páginas de la matriz, se encontraron otros tipos de página:

4. Páginas en las cuales se tiene el nombre de la materia, pero no hay información de algún grupo, por ejemplo <http://www.fciencias.unam.mx/docencia/horarios/20081/1556/803>



5. Páginas que no tienen información del salón: <http://www.fciencias.unam.mx/docencia/horarios/20081/119/4>

Actuaría (plan 2000)

Álgebra Moderna IV, Optativas

Grupo 4250 6 alumnos.

Profesor José Ríos Montes lu mi vi 13 a 14

Ayudante

ma ju 13 a 14

6. Páginas que tienen grupos sin información del número de alumnos:
<http://www.fciencias.unam.mx/docencia/horarios/20112/119/630>

Actuaría (plan 2000)

Procesos Estocásticos I, Optativas

Grupo 6157 2 alumnos.

Profesor Fernando Guerrero Poblete lu mi vi 12 a 13 O216

Ayudante Héctor Alonso Olivares Aguayo ma ju 12 a 13 O216

Ayudante Rafael Martínez Sánchez ma ju 12 a 13

Ayudante Alfredo Hernández Lammoglia ma ju 12 a 13

Grupo 6192 3 alumnos.

Profesor Guillermo Garro Gómez lu mi vi 18 a 19 O122

Ayudante Martín Martínez Estrada ma ju 18 a 19 O122

Grupo 6193

Profesor Fernando Baltazar Larios lu mi vi 17 a 18 O221

Ayudante Estela Eréndira Zamora García ma ju 17 a 18 O221

7. Páginas que tienen grupos sólo con el horario (sin nombre del profesor, salón, ayudante, número de alumnos, lugares disponibles):
<http://www.fciencias.unam.mx/docencia/horarios/20091/119/841> <http://www.fciencias.unam.mx/docencia/horarios/20091/119/244>

Actuaría (plan 2000)

Variable Compleja II, Optativas

Grupo 4521 33 alumnos.

Profesor Guillermo Javier Francisco Sienra Loera lu mi vi 12 a 13 O123

Ayudante Adriana Andraca Gómez ma ju 12 a 13 O123

Presentación

Grupo 4519

Profesor lu mi vi 17 a 18

Ayudante ma ju 17 a 18

- Una vez que se probó la función `posibles_url` se obtuvo una primera matriz llamada `mat_posibles_url`, con 18725 entradas, ahí se observó que el número máximo de materia es 991, por lo que el valor máximo se puede reducir de 10000 a 1000.
- La matriz `mat_posibles_url` se define con un tamaño fijo antes de correr el algoritmo para que no se demore por tener un objeto que va cambiando de tamaño, por lo que al final de haberle aplicado la función se le deben de quitar los renglones que no tienen información.
- Dentro de los problemas de información repetida, se encontraron 2 posibles casos:
 1. Tener información de una materia correspondiente a un plan de estudios posterior al semestre: <http://www.fciencias.unam.mx/docencia/horarios/20082/1556/803> y tener la misma información con el plan de estudios correspondiente: <http://www.fciencias.unam.mx/docencia/horarios/20082/218/803>

Ciencias de la Computación (plan 1994)

Graficación por Computadoras, Optativas

Grupo 7054 19 alumnos.

Profesor Ana Luisa Solís González-Cosío lu mi vi 12 a 13

Ayudante José Israel Figueroa Angulo ma ju 12 a 13

Ayud. Lab. Azael Nieves Ramírez

Ciencias de la Computación (plan 2013)

Graficación por Computadoras, Optativas

Grupo 7054 19 alumnos.

Profesor Ana Luisa Solís González-Cosío lu mi vi 12 a 13

Ayudante José Israel Figueroa Angulo ma ju 12 a 13

Ayud. Lab. Azael Nieves Ramírez

2. Tener una misma materia con nombres distintos para las diferentes carreras: <http://www.fciencias.unam.mx/docencia/horarios/20201/2017/1739> para Actuaría, plan 2015 y <http://www.fciencias.unam.mx/docencia/horarios/20201/217/1712> para Matemáticas, plan 1983. Notamos que la información en ambas páginas es la misma, sólo se cambia el orden en la que se encuentran los grupos.

Matemáticas (plan 1983)

Estadística III, Optativas de los Niveles VII y VIII

Grupo 9259, 35 lugares, 11 alumnos.

Profesor Claudia Lara Pérez Soto lu mi vi 9 a 10 101 (Nuevo Edificio)

Ayudante Ventura Jimenez Martinez ma ju 9 a 10 101 (Nuevo Edificio)

Grupo 9261, 81 lugares, 32 alumnos.

Profesor Sofia Villers Gómez lu mi vi 9 a 10 306 (Yelizcalli)

Ayudante Amílcar José Escobedo Pérez ma ju 9 a 10 306 (Yelizcalli)

Grupo 9263, 56 lugares, 9 alumnos.

Profesor Luis Antonio Rincón Solís lu mi vi 9 a 10 P102

Ayudante José Luis Miranda Olvera ma ju 9 a 10 P102

Actuaría (plan 2015)

Modelos de Supervivencia y de Series de Tiempo, Séptimo Semestre

Grupo 9258, 35 lugares, 11 alumnos.

Profesor Claudia Lara Pérez Soto lu mi vi 9 a 10 101 (Nuevo Edificio)

Ayudante Ventura Jimenez Martinez ma ju 9 a 10 101 (Nuevo Edificio)

Grupo 9260, 81 lugares, 32 alumnos.

Profesor Sofia Villers Gómez lu mi vi 9 a 10 306 (Yelizcalli)

Ayudante Amílcar José Escobedo Pérez ma ju 9 a 10 306 (Yelizcalli)

Grupo 9262, 56 lugares, 9 alumnos.

Profesor Luis Antonio Rincón Solís lu mi vi 9 a 10 P102

Ayudante José Luis Miranda Olvera ma ju 9 a 10 P102

- En la función `genera_m13` se obtiene una matriz con 13 columnas: Materia, Profesor, Horario, Lugares, Alumnos, Salón, Grupo, Carrera, Plan, Semestre, Cambios, Turno, Sem. Materia.

La columna 11, *Cambios*, va a guardar todos los cambios que han “sobrevivido” esas páginas, dependiendo los números que se encuentren serán los cambios de cada url.

- (1) Corresponde a que se eliminaron las copias de los planes de estudio viejos
- (2) Corresponde a que se eliminaron las copias de las materias que tienen nombres distintos para carreras diferentes
- (3) Corresponde a las páginas que no tienen información del salón

La columna 12, *Turno*, va a ser una columna binaria en la cual habrá un 1 si el turno es matutino el cual consideraremos que será desde las 7am hasta las 2pm, tomando en cuenta la clase que comienza a las 2pm y termina a las 3pm; y habrá un 0 si el turno es vespertino, considerado de 3pm a 9pm, contando como clase inicial la de 3 – 4pm y como la última clase la de 9 – 10pm

La columna 13, *Sem. Materia*, contiene el semestre en el que el plan de estudios dicta que se lleva esa materia. Ej. “Primer”, “Segundo”, ... “Octavo”, “Optativas”

- Las categorías de las columnas de *m5* son:

Col.	Nombre	Explicación	Posibles
1	Semestre	Semestre al que pertenece la materia (Año y semestre)	1º, 2º, ..., 8º
2	Plan	Año en el que se implemento un nuevo plan de estudios	Número de plan de estudios
3	Materia	Clave del curso impartido	\mathbb{N}
4	URL	Nombres de las páginas de los horarios de FC	Nombres de páginas de internet
5	Num. Grupos	Número de grupos que hay en cada página de internet	\mathbb{N}

Semestre	Plan	Materia	URL
20182	2017	92	http://www.fciencias.unam.mx/docencia/horarios/20182/2017/...
20192	2017	92	http://www.fciencias.unam.mx/docencia/horarios/20192/2017/...
20151	2017	625	http://www.fciencias.unam.mx/docencia/horarios/20151/2017/...
20161	2017	625	http://www.fciencias.unam.mx/docencia/horarios/20161/2017/...

- Las categorías de las columnas de $m13$ son:

Col.	Nombre	Explicación	Posibles valores
1	Materia	Nombre del curso impartido	Nombres de materias
2	Profesor	Nombre de la persona que va a impartir alguna materia	Nombres de personas
3	Horario	Hora en la que se imparte alguna materia	15 horas en total
4	Lugares	Espacios disponibles por salón	\mathbb{N}
5	Alumnos	Número de estudiantes inscritos por grupo	\mathbb{N}
6	Salón	Espacio físico en el que se imparte alguna materia	Nombres de salones
7	Grupo	Clave con la que se identifica una asignación	Clave de grupos
8	Carrera	Nombre de alguna carrera de FC	Actuaría, C. de la C., Mate, Mate Ap.
9	Plan	Año en el que se implemento un nuevo plan de estudios	Número de plan de estudios
10	Semestre	Semestre al que pertenece la materia (Año y semestre)	1º, 2º, ..., 8º
11	Cambios	Clave que indica los cambios que se le han hecho a la materia	1, 2, 3
12	Turno	Matutino: 7:00-14:00hrs, Vespertino: 15:00-21:00	Matutino/Vespertino
13	Sem.Materia	Semestre en el que el plan de estudios dicta que se lleva esa materia	“Primer”, “Segundo”, ... “Octavo”, “Optativas”

- Se obtuvieron 4 grupos de datos para hacer los análisis estadísticos:

G_1 : Turno matutino semestre impar

G_2 : Turno vespertino semestre impar

G_3 : Turno matutino semestre par

G_4 : Turno vespertino semestre par

Sem. \ Turno	Matutino	Vespertino
Impar	G_1	G_2
Par	G_3	G_4

- Problemas con la extracción de información en las páginas web

1. Dentro de la obtención de datos del número de alumnos, no se lee la información cuando se tiene *Un alumno*