

REpte BIG DATA

És el Leonardo un institut segur?



Júlia Botey
Núria Graner
Noel Corchón
Pau Torrellas

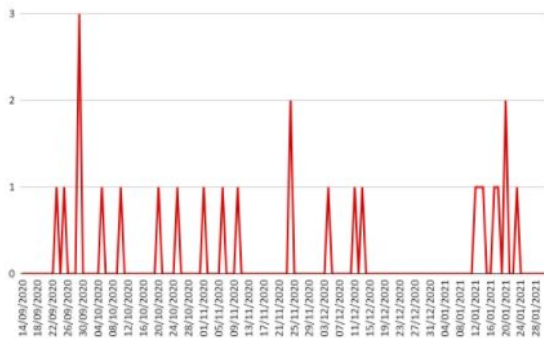
Institut Leonardo da Vinci
Sant Cugat del Vallès

És el Leonardo un institut segur?

L'institut Leonardo da Vinci es troba ubicat al municipi de Sant Cugat del Vallès i imparteix classes d'Educació Secundària Obligatòria i Batxillerat. En el període lectiu entre l'inici de curs, el 14 de setembre de 2020 i el 31 de gener de 2021 a l'institut hem tingut un total de 25 casos positius de Covid-19. I ens plantegem si 25 positius és poc o molt. És, el nostre centre, més o menys segur que la resta d'instituts de Catalunya?

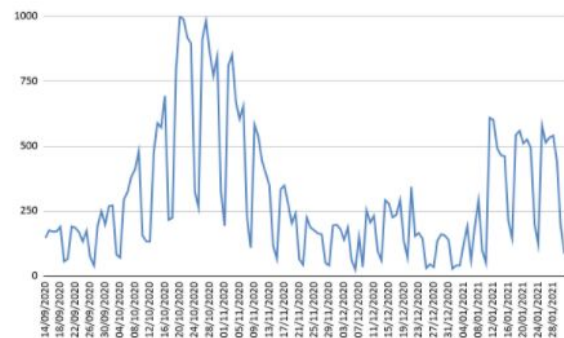
A continuació, analitzem les dades obtingudes en el nostre centre i les comparem amb les dades obertes que es poden consultar a <https://analisi.transparenciacatalunya.cat/>, la pàgina de referència del Departament de Salut de la Generalitat de Catalunya.

Institut



3,5%

Catalunya

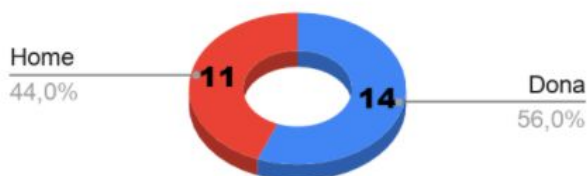


5,06%

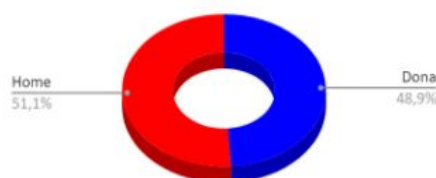
El 3,5% dels alumnes de l'institut han estat contagiats, mentre que les dades de l'esmentada base de dades relatives a joves entre 10 i 19 anys de Catalunya indiquen que s'ha contagiats un 5,06%; per tant a l'institut hem tingut un 30,9% menys de contagis. A més en els gràfics podem veure que el màxim de persones contagiades en un dia a l'institut són 3 (per tres persones de batxillerat). A tot Catalunya el màxim en un dia - de la franja d'edat entre 10 i 19 anys - són 1000 persones contagiades.

Una dada curiosa és que entre els alumnes contagiats de l'institut hi ha més **dones** contagiades (14 persones, 56%) que **homes** (11 persones, 44%) quan realment les dones són el 45.4% del total dels alumnes del centre. Però a Catalunya passa al revés: hi ha menys contagis entre **dones** (48%) que **homes** (51%), i correspon exactament al percentatge de **dones** de la població (48.1%) i d'**homes** (51.9%) d'aquesta edat:

Casos positius per home/dona a l'institut

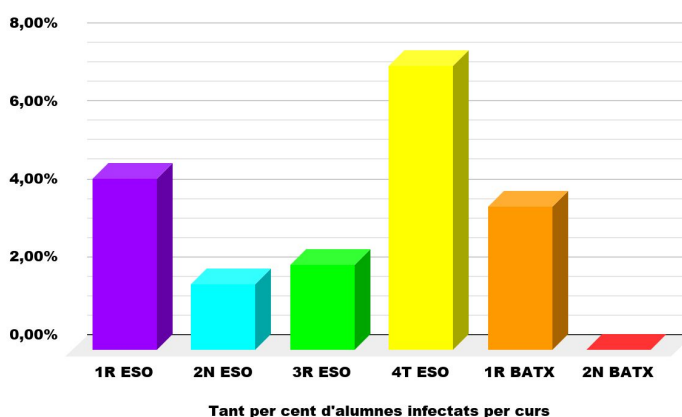


CASOS POSITIUS PER HOMES/DONES GENERALITAT DE CATALUNYA



Fins ara hem estat comparant les dades de Catalunya amb les del nostre l'institut però anem a veure què passa amb els 25 contagiats del nostre centre: del primer que ens n'adonem és que a l'institut, s'han contagiats més els alumnes d'uns cursos que dels altres:

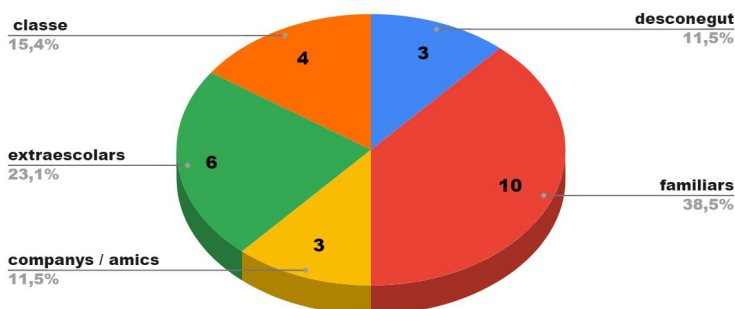
En el gràfic es pot veure que a **2n de batxillerat**, no s'ha contagiats ningú, cal destacar que només hi ha 25 alumnes en aquest curs. A **1r de batxillerat** s'ha contagiats el 3,7% dels alumnes. Tot i que el batxillerats han fet la meitat de les classes online. En canvi, a **4t d'ESO**, què és on s'ha contagiats més gent, (el 7,3%) hi ha menys alumnes que a **1r**, **2n** o **3r** d'ESO.



1r d'eso ha sigut el segon grup més infectat, amb un 4,4% d'alumnes contagiats. Entre els alumnes contagiats de 4t i 1r d'ESO fan més de la meitat de contagis. Tot i les dades que tenim, no hem pogut esbrinar perquè uns nivells es contagien més que d'altres.

Ja sabent a quin nivell anaven els contagiats, ara veurem on es van contagiar. Segons les fonts pròpies de l'institut:

On s'ha contagiats?



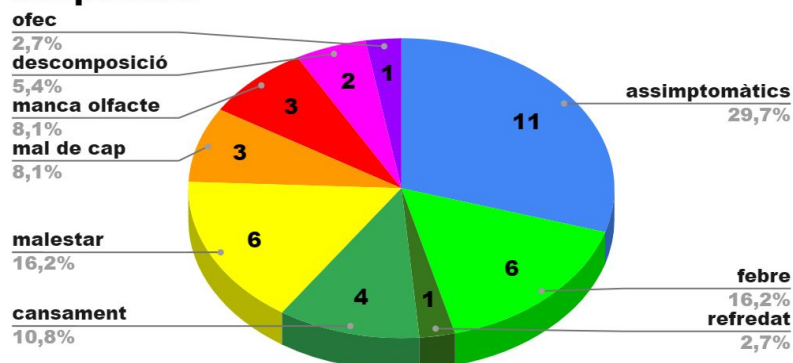
Al contrari del que molta gent pensa, l'institut no és el principal focus de contagi, ja que a **classe** només s'han contagiats 4 alumnes, un 15% dels contagiats. El principal grup infeccions són els **familiars**, un total de 10 alumnes que són el 38% dels contagiats, han estat contagiats a casa o en reunions

familiars. A les **extraescolars** s'han contagiats 6 alumnes, que representen el 23%. És el segon motiu de contagi, després de les reunions familiars. 3 dels casos del Leonardo són d'origen **desconegut**, ja que no s'ha pogut esbrinar on s'ha infectat l'alumne en qüestió. Per tant, la conclusió és que els majors focus de contagi són les activitats extraescolars, i els familiars, i per tant són els àmbits on s'hauria de fer més prevenció.

Un altre tema a observar és quins símptomes han tingut els alumnes contagiats. Aquí tenim un gràfic elaborat a partir de les dades de l'institut:

La majoria d'alumnes han sigut **assimptomàtics**, amb 11 alumnes, el 29,7%. El símptomes més comuns han sigut la **febre** i el **malestar** ja que sis persones han patit aquests símptomes. Només una persona ha patit **ofec** i només una símptomes de **refredat**. La resta de símptomes estan repartits més homogeneament.

Símptomes



Per últim, i com a conclusió, podem dir que l'Institut Leonardo da Vinci és un institut segur amb només 25 casos, que són molt pocs, ja que, comparant les dades, observem que a Catalunya s'ha contagiats el 5,06% de la població, en canvi, a l'institut només s'ha contagiats el 3,5%. Per tant, el Leonardo és un institut segur que està per sota de la mitjana de contagis.

FONTS CONSULTADES:

<http://educacio.gencat.cat/ca/actualitat/escolasegura/tracacovid/dades-covid19-centres/#/>

<https://analisi.transparenciacatalunya.cat/d/qwj8-xpvk/visualization>

Dades pròpies del centre, facilitades per Anabel Esteban, Coordinadora Covid del centre.

<https://www.naciodigital.cat/noticia/210555/coronavirus-mapes-grafics-Catalunya-Covid-dades-evolucio-pandemia>

<http://cbiolegs.cat/wp-content/uploads/2021/02/FAQsvirusA4.pdf>

https://elpais.com/educacion/2021-02-10/un-alumno-infectado-contagia-a-04-personas-en-la-escuela.html?prm=enviar_email

<https://elpais.com/ciencia/2020-10-24/un-salon-un-bar-y-una-clase-asi-contagia-el-coronavirus-en-el-aire.html>

Les cinc V del Big Data

Primera V *volum*

Hem comparat les dades obertes corresponents al període escolar comprès entre el 14 de setembre de 2020 i el 31 de gener de 2021 dels alumnes d'entre 10-19 anys de Catalunya amb les del nostre centre.

Per gestionar aquest volum de dades hem utilitzat Excel i ens hem ajudat de gràfics, més comprensibles per a tots.

Les dades de l'institut son massa poques com per extreure conclusions extrapolables a poblacions més grans. Potser si haguéssim pogut acumular les dades de tot el curs podríem haver arribat a conclusions més acurades.

Segona V *velocitat*

Pensem que aquestes dades poden seguir essent vàlides en estudis posteriors que vulguin investigar aquesta pandèmia si són certes, cosa amb la que hem treballat i en la que confiem.

Creiem que, si haguéssim pogut crear abans de començar el curs una pàgina on registrar símptomes, data de PCR, origen del contagi...Hauria estat molt més ràpid recollir les dades. I si haguéssim pogut replicar un model general en dades obertes, les comparacions haurien estat més exactes

Tercera V *varietat*

Les nostres dades provenen de la base de dades de l'institut i de les dades obertes de Catalunya, extretes del Portal de Dades Obertes de Catalunya. Ens hauria agradat, per exemple, completar l'anàlisi amb dades d'altres instituts de Sant Cugat. També hauria estat millor poder obtenir la mateixa franja exacta d'edat, doncs hem hagut de comparar la franja de 10 a 19 anys que apareix al Portal de Dades, amb les edats de l'institut, entre 11 i 18 anys.

Quarta V *veracitat*

El que no podem controlar és si hi ha persones que puguin haver donat positiu per Covid-19 i no haver informat al centre, per exemple, entre l'alumnat de Batxillerat, que només ha fet la meitat de les classes presencials durant una part del temps analitzat. Tampoc sabem si poden haver hagut casos de persones asimptomàtiques que hagin contagiats altres asimptomàtics i en no tenir ningú símptomes, no hagin estat detectats i per tant són casos no contabilitzats.

Cinquena V *visualització*

Ho hem fet visible amb gràfics fets a partir les dades de l'institut i de Catalunya ficades en un full de càlcul, que permet convertir les dades en gràfics i és més visual, i després les hem comparat.