

# Corona-Zusatzbefragung zur Volkshochschul-Statistik – DIECovidSurvey

*Methodenbericht*

Version 1.0.0, 11.10.2022

Dr. Kerstin Hoenig

Deutsches Institut für Erwachsenenbildung

Leibniz-Zentrum für Lebenslanges Lernen e. V.

## Konzept

Das DIECovidSurvey wurde vom Deutschen Institut für Erwachsenenbildung (DIE) in Zusammenarbeit mit dem Deutschen Volkshochschulverband (dvv) im Herbst 2020 durchgeführt, um die Auswirkungen der Corona-Pandemie auf die deutschen Volkshochschulen (vhs) zu untersuchen. Die Daten können mit den jährlich vom DIE erhobenen Daten der Volkshochschul-Statistik kombiniert werden, die als Scientific Use File zur Verfügung stehen (Echarti et al, 2022; Deutsches Institut für Erwachsenenbildung, 2022).

Der Fragebogen wurde von DIE und dvv gemeinsam entwickelt. Kernstück der Befragung sind detaillierte Angaben über das Veranstaltungsangebot im ersten Lockdown im Frühling und Frühsommer 2020, als Veranstaltungen in Präsenz untersagt waren. Der Fragebogen erhebt detailliert für jeden Programmbereich die Zahl der vor Lockdown geplanten sowie im Lockdown tatsächlich durchgeführten Kurse und Einzelveranstaltungen sowie das Veranstaltungsformat (Präsenz/Blended Learning/Online). Weitere Inhalte der Befragung betreffen die personelle und finanzielle Situation, das verfügbare Raumangebot, Auswirkungen der Pandemie auf Teilnehmendengruppen, die Nutzung digitaler Technologien inklusive der vhs.cloud, die Aufnahme coronabezogener Veranstaltungen ins Programm, Einschätzungen der Lage zum Befragungszeitpunkt, sowie zukünftige Strategien und wahrgenommene Herausforderungen in Bezug auf Digitalisierung und Programmgestaltung.

Die Erhebung wurde als Online-Befragung in LimeSurvey durchgeführt, wobei eine Aufforderung zur Teilnahme an alle deutschen vhs erging.

## Datenerhebung und Stichprobe

### *Befragungsinstrument*

Der CAWI-Fragebogen wurde in LimeSurvey programmiert und anschließend von Mitarbeitenden von DIE und dvv getestet. Die Umfrage bestand aus 70 Frageblöcken mit insgesamt 317 Items. Davon entfallen 160 Items auf die quantitative Erfassung der für den Zeitraum geplanten und tatsächlich durchgeführten Kurse und Einzelveranstaltungen nach Programmbereich und Veranstaltungsformat.

Bis auf wenige Eingangsisems, deren Angaben für die Filterführung der Befragung erforderlich waren, war die Beantwortung der Items nicht verpflichtend. Quantitative Angaben zu Veranstaltungen, Raumkapazitäten etc. wurden in der Programmierung der Befragung auf interne Plausibilität geprüft und zur Eingabehilfe und -prüfung Berechnungstabellen mit den gemachten Angaben und Summenwerten ausgespielt. Im Falle unplausibler oder fehlender Angaben (z. B., wenn die Summe der weitergeführten die Summe der geplanten Veranstaltungen überstieg) erschienen Warnhinweise. Eine Korrektur der gemachten Eingaben wurde jedoch nicht erzwungen.

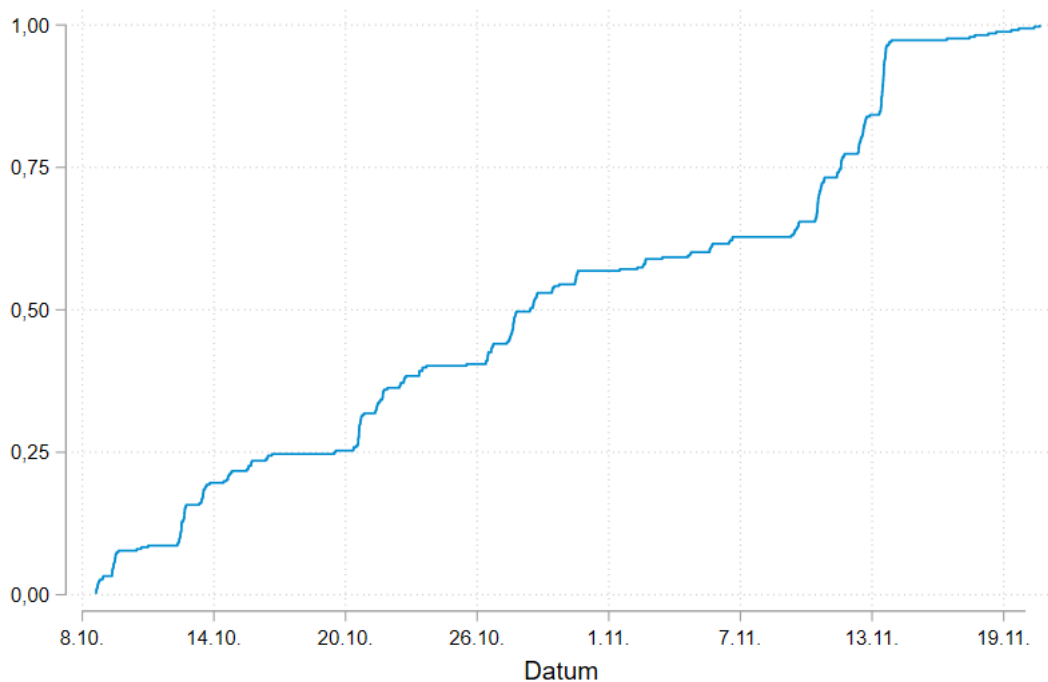
### *Einsatzstichprobe und Kontaktierung*

Die Einsatzstichprobe bestand aus der Grundgesamtheit aller deutschen vhs (N=870 im Jahr 2020). Aufgrund der jährlichen Datenerhebung im Rahmen der

Volkshochschul-Statistik bestehen an allen deutschen vhs Ansprechpartner für die Statistik, die vom DIE am 8.10.2020 per E-Mail kontaktiert wurden mit Bitte um Teilnahme und einem Link zum Online-Fragebogen. Ebenfalls wurde den vhs zur Vorbereitung der Fragebogen im PDF-Format zugeschickt. Der Befragungslink war nicht personalisiert, stattdessen erfolgte die Identifikation der vhs über Eingabe der Postleitzahl und der aus der Volkshochschul-Statistik bekannten VHS-ID zu Beginn der Befragung. Im Vorfeld waren die Landesverbände der Volkshochschulen durch den dvv gebeten worden, ihre vhs über die bevorstehende Befragung zu informieren und um Teilnahme zu bitten. Ein Reminder zur Teilnahme wurden per E-Mail am 26.10. durch das DIE verschickt.

#### *Feldphase und Bearbeitungsdauer*

Die Feldphase war ursprünglich vom 08.10.-30.10.2020 geplant. Aufgrund von Rückmeldungen aus dem Feld wurde diese bis zum 20.11.2020 verlängert. Die untenstehende Grafik zeigt den Rücklauf über die Zeit.



*Abbildung 1: Rücklauf über die Feldzeit*

Laut Rückmeldungen aus dem Feld benötigten viele vhs mehrere Tage, um die benötigten Informationen zu sammeln. Dies drückt sich jedoch nicht immer in der in Limesurvey erfassten Bearbeitungsdauer, gemessen als Beginn der Bearbeitung der Umfrage in Limesurvey und der letzten Eingabe, aus. Diese variiert zwischen den teilnehmenden vhs stark, da einige vhs im Vorfeld anhand des PDF-Befragungsbogens die benötigten Informationen sammelten und dann in Limesurvey eintrugen, während andere die Daten schrittweise beim Durchklicken der Befragung in Limesurvey recherchierten und eingaben. Die vhs konnten Befragungsdaten zwischenspeichern und die Umfrage zu einem späteren Zeitpunkt fortsetzen.

### *Rücklauf und Ausschöpfungsquote*

Da die Programmplanung an größeren vhs arbeitsteilig von den zuständigen Programmbereichsleitenden übernommen wird, wurde der Fragebogen an einigen vhs mehrfach von den zuständigen Programmbereichsleitenden jeweils für ihren Programmbereich ausgefüllt. Diese Fälle wurden bei der Datenaufbereitung harmonisiert (s.u.). Die folgenden Tabellen geben Auskunft über den Rücklauf:

*Tabelle 1: Befragungsteilnahmen*

Brutto Öffnungen Umfrage (Klick auf Link)	1015
Davon ohne ausgefüllte Filtereingangsfragen	604
VHS nicht identifizierbar über VHS-ID & PLZ	8
<b>Zwischensumme gültige Beobachtungen</b>	<b>403</b>
davon Duplikate (Umfrage von mehreren Beteiligten an einer vhs ausgefüllt)	67
<b>Harmonisierte gültige Beobachtungen (1 Beobachtung pro vhs)</b>	<b>336</b>
<b>Davon Umfrage vollständig bearbeitet und abgesendet</b>	<b>233</b>

*Tabelle 2: Übersicht über mehrfach ausgefüllte Befragungsbögen pro vhs*

<b>Ausgefüllte Befragungen pro vhs</b>	<b>N Antworten</b>	<b>N vhs</b>
1	283	283
2	82	41
3	33	11
5	5	2
<b>Gesamt</b>	<b>403</b>	<b>336</b>

Ignoriert man testweise Öffnungen der Umfrage ohne gültige Eingaben auf die verpflichtenden Filtereingangsfragen, auf denen die restliche Befragung beruht, so bleiben 411 Fälle, in denen mit einer Bearbeitung des Fragebogens zumindest begonnen wurde. Von diesen sind 8 nicht identifizierbar, da keine gültige VHS-ID oder PLZ angegeben wurde (s. Datenaufbereitung). Damit bleiben 403 Fälle, die eindeutig einer vhs zuordbar sind. Nach Harmonisierung der Mehrfachantworten innerhalb derselben vhs (vgl. Tabelle 2) verbleiben 336 Fälle. Dies entspricht 39% aller vhs. Von diesen haben 233 Fälle (27%) die Umfrage vollständig bearbeitet und abgeschickt.

### *Kontaktaufnahmen und Rückmeldungen zur Befragung*

Während der gesamten Feldphase bestand die Möglichkeit, telefonisch und per Email Kontakt mit den zuständigen Mitarbeitenden am DIE sowie mit einer Ansprechperson des dvv aufzunehmen. Die häufigsten Rückmeldungen betrafen die Bitte, die Befragungsdauer zu verlängern, da in einigen Bundesländern durch die Herbstferien viele Beschäftigte an den vhs in Urlaub waren. Weitere Anfragegründe waren technischer Natur (Probleme bei der Zwischenspeicherung oder Öffnung der gespeicherten Daten; versehentliches Absenden der Daten vor tatsächlicher Fertigstellung der Befragung; etc.) sowie die Bitte, die ausgefüllten Befragungsdaten tabellarisch zugesandt zu bekommen, da die vhs die zusammengetragenen

Informationen teilweise selbst nicht systematisch abgelegt hatten. Diesem Wunsch wurde durch Zusendung einer Excel-Tabelle der für die jeweilige vhs eingetragenen Daten entsprochen. Von einigen vhs wurde die Kontaktaufnahme genutzt, um Unmut über die Dauer und Detailliertheit der Befragung zu äußern oder sich für die Nichtteilnahme aufgrund hoher Arbeitsbelastung zu entschuldigen. Es gab keine Verständnisfragen zu den Inhalten der Befragung.

## Datenaufbereitung

### *Datenerfassung*

Die Befragungsdaten wurden von Limesurvey in Stata (Version 17.0) importiert und dort aufbereitet. Folgende Bearbeitungsschritte wurden durchgeführt:

1. **Löschung von testweisen Öffnungen der Befragung:** Da die Befragung über einen nicht personalisierten Befragungslink zugänglich war, wurde der Fragebogen oft testweise geöffnet und dann wieder beendet. Beobachtungen, die die verpflichtenden Filtereingangsfragen (VHS-ID und PLZ sowie die Variablen progbeg, progend, panbegyn, panendyn, onlineyn) zu Beginn der Befragung nicht vollständig beantwortet haben, wurden gelöscht. Ohne Vervollständigung dieser Angaben war ein weiteres Bearbeiten der Befragung nicht möglich, da auf ihnen die Filterung beruht. Dies betraf 604 von 1015 Aufrufen der Befragung.
2. **Identifikation der vhs anhand von VHS-ID und PLZ:** Für alle Eingaben wurde kontrolliert, ob die eingegebene PLZ und VHS-ID übereinstimmen mit einer vhs aus der Adressliste des DIE. Wenn nur eine der beiden Eingaben mit der Adressliste übereinstimmte, wurde die eingegebene PLZ oder VHS-ID auf Fehler überprüft (Vertipper, Zahlendreher, PLZ einer vhs-Nebenstelle). In 8 Fällen mit Dateneingabe konnte über Abgleich von VHS-ID und PLZ keine vhs eindeutig identifiziert werden. Eine Prüfung der eingegebenen Werte legt die Vermutung nahe, dass es sich hierbei um nicht valide Testeinträge handelt, da ohne Eingabe von fünfstelligen Werten für VHS-ID und PLZ der Fragebogen nicht zugänglich war. Diese Fälle wurden gelöscht.
3. **Harmonisierung von doppelt beantworteten Fragebögen zu einer Beobachtung:** Mehrfachantworten aus einer vhs wurden zu einer Beobachtung zusammengefasst, indem die Antworten sortiert wurden nach Vollständigkeit (Fragebogen abgeschickt ja/nein) und Datum (absteigend). Pro Variable wurde jeweils die erste gültige Antwort übernommen. Dadurch kann es in Einzelfällen vorkommen, dass die harmonisierten Datenzeilen Einträge enthalten, die der Filterführung der Befragung widersprechen. Dem Datensatz wurde eine Variable hinzugefügt, die harmonisierte Beobachtungen kennzeichnet.
4. **Anonymisierungsmaßnahmen:** Um eine Reidentifizierung einzelner vhs zu verhindern wurden die Postleitzahlen gelöscht und die VHS-IDs gegen die Scientific Use File IDs der Volkshochschul-Statistik getauscht. Über diese IDs ist auch eine Kombination der beiden Datensätze möglich. Nutzende mit entsprechender Berechtigung können über das DIE zusätzliche regionale

Informationen (z. B. Bundesland, Kreiskennziffer) beziehen. Alle offenen Angaben wurden geprüft auf Nennung persönlicher Daten oder Begriffen, die eine Identifikation von vhs zulassen oder Hinweise auf das betreffende Bundesland enthalten und ggf. anonymisiert durch Ersetzung der betreffenden Textstellen durch den Hinweis „[anonymisiert]“.

- 5. weitere Datenaufbereitungen, die der besseren Lesbarkeit und Weiterverarbeitung der Daten dienen.** Dies betrifft die Formatierung von Datumsangaben, Konvertierung von Stringvariablen in numerische Variablen (soweit möglich) und Kürzung von Labels und Variablennamen. Allen Variablen wurde ein Hinweis (note-Befehl in Stata) mit Verweis auf die Fragennummer hinzugefügt zum besseren Abgleich von Fragebogen, Codebuch und Datensatz.

### *Datenprüfung*

Bei Durchsicht der Daten ist ersichtlich, dass viele vhs die Umfrage nicht vollständig bearbeitet haben. So haben von 336 vhs, die die Befragung begonnen haben, nur 233 die Befragung über den „Beenden“-Button offiziell beendet. Da die Beantwortung der meisten Items nicht verpflichtend war, sind auch unter diesen 233 Beobachtungen Fälle enthalten, die nur Teile der Befragung beantwortet haben. Tabelle 3 stellt den Anteil fehlender Werte, die nicht auf die Filterführung zurückzuführen sind, nach Beendigungsstatus dar. Teilweise sind hohe Missing-Raten dadurch zu erklären, dass einige vhs bei den quantitativen Angaben zu Raumangebot und Veranstaltungsprogramm entgegen den Hinweisen im Fragebogen keine Nullen eingegeben haben.

*Tabelle 3: Anteil nicht beantworteter Items nach Beendigungsstatus (in Prozent)*

	<b>Mean</b>	<b>SD</b>	<b>Median</b>	<b>P10</b>	<b>P90</b>
<b>abgebrochen (N=103)</b>	79,6	19,1	88,3	60,0	92,1
<b>beendet (N=233)</b>	12,0	16,7	5,7	3,2	31,7
<b>Gesamt (N=336)</b>	32,7	35,7	7,9	3,5	88,9

Zur Prüfung der Repräsentativität wurden die Daten aus dem Berichtsjahr 2019 der Volkshochschul-Statistik verwendet. Dabei handelt es sich um eine Vollerhebung aller deutschen Volkshochschulen mit einer Beteiligungsrate von 99% (Huntemann et al., 2021).

Mit Ausnahme von Hamburg haben sich vhs aus allen Bundesländern beteiligt. Besonders hoch war die Teilnahmebereitschaft im Saarland (62,5%), in Bremen (50%) und Mecklenburg-Vorpommern (50%). Die geringsten Teilnahmequoten wurden in Schleswig-Holstein (23%) und Berlin (25%) verzeichnet (vgl.

). Die Teilnahmebereitschaft fiel im ländlichen Raum etwas geringer aus als in Städten und Großstädten (vgl.

Tabelle 5).

Als Maß für die Größe der vhs wird die Summe der Unterrichtsstunden im Jahr 2019 verwendet. Wie in Abbildung 2 zu sehen ist, sind kleine vhs mit wenig Unterrichtsstunden pro Jahr im Sample unterrepräsentiert.

Tabelle 4: Teilnahme nach Bundesland

Bundesland	nicht teilgenommen	teilgenommen	Gesamt
Schleswig-Holstein	77,0	23,0	100,0
Hamburg	100,0	0,0	100,0
Niedersachsen	54,4	45,6	100,0
Bremen	50,0	50,0	100,0
NRW	52,0	48,0	100,0
Hessen	56,3	43,8	100,0
Rheinland-Pfalz	56,1	43,9	100,0
Baden-Württemberg	59,2	40,8	100,0
Bayern	63,3	36,7	100,0
Saarland	37,5	62,5	100,0
Berlin	75,0	25,0	100,0
Brandenburg	73,7	26,3	100,0
Mecklenburg-Vorpommern	50,0	50,0	100,0
Sachsen	68,8	31,3	100,0
Sachsen-Anhalt	66,7	33,3	100,0
Thüringen	69,6	30,4	100,0
<b>Gesamt</b>	<b>61,8</b>	<b>38,2</b>	<b>100,0</b>
Cramérs V	0,19	Chi <sup>2</sup> (p)	30,3 (0,0109)

Tabelle 5: Teilnahme nach Kreistyp

	Kreisfreie Großstädte	städtische Kreise	Ländl. Kreise mit Verdich- tungsan- sätzen	Dünn besiedelte ländliche Kreise	Gesamt
nicht teilgenommen	59,7	56,3	70,4	65,7	62,0
teilgenommen	40,3	43,7	29,6	34,3	38,0
<b>Gesamt</b>	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>
Cramérs V	0,12		Chi <sup>2</sup> (p)	13,32 (0,004)	
Gamma (ASE)	-0,16 (0,055)		Tau b (ASE)	-0,09 (0,031)	

Betrachtet man die Selektivität nach Erfahrung mit digitalen Lernangeboten vor Pandemiebeginn, so ist der Anteil von vhs, die bereits 2019 digitale Lernangebote im Programm hatten, unter den teilnehmenden vhs mit 49,6% etwas höher als unter den nicht teilnehmenden mit 43,3%. Die Differenz ist jedoch nicht statistisch signifikant auf dem 5%-Niveau ((prop(nicht teilgenommen) - prop(teilgenommen)) = -0,061; p = 0,075).

Aufgrund der Möglichkeit, Daten der Volkshochschul-Statistik anzuspielen, ist die Konstruktion von Gewichten oder die Anwendung (multipler) Imputationsverfahren für fehlende Werte/Beobachtungen grundsätzlich möglich.

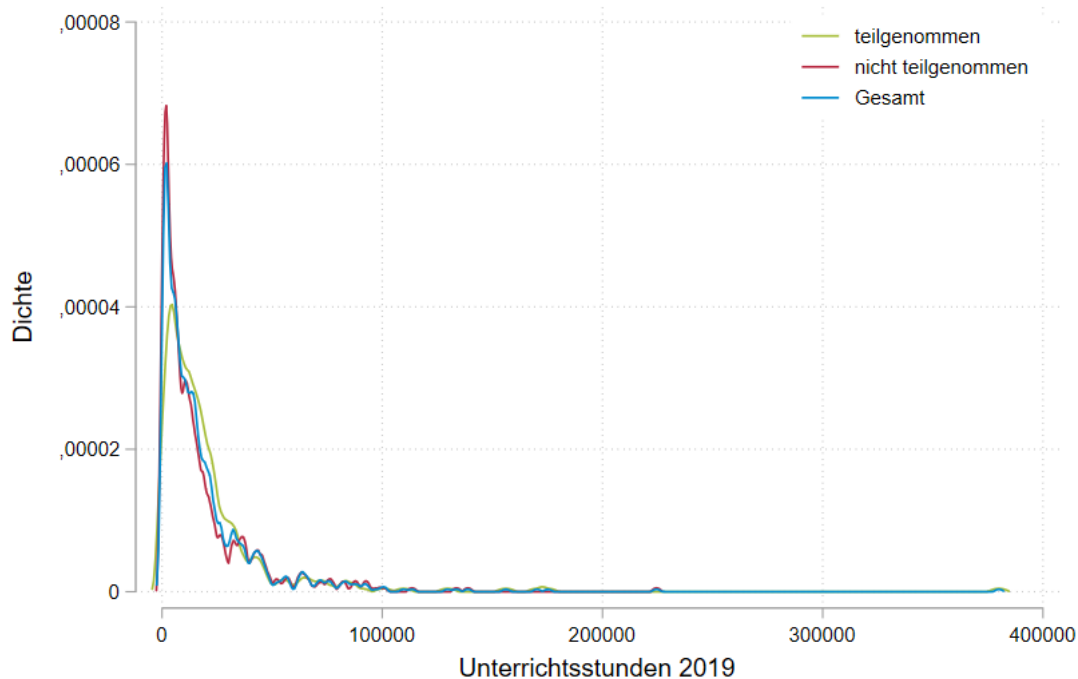


Abbildung 2: Unterrichtsstunden im Jahr 2019 nach Teilnahmestatus  
(Epanechnikov-Kerndichteschätzer, Sheather-Jones Bandwidth Selector)

## Literatur

Deutsches Institut für Erwachsenenbildung (2022). *Scientific Use File Volkshochschul-Statistik* [Datensatz]. GESIS, Köln. ZA6276 Datenfile Version 1.0.0. <https://doi.org/10.4232/1.13953>

Echarti, N., Huntemann, H., Reichart, E. & Lux, T. (2022). *Volkshochschul-Statistik: 59. Folge, Berichtsjahr 2020*. wbv Media. <https://doi.org/10.3278/I70681>

Huntemann, H., Echarti, N., Lux, T. & Reichart, E. (2021). *Volkshochschul-Statistik: 58. Folge, Berichtsjahr 2019*. wbv Media. <https://doi.org/10.3278/70681>